DOCUMENT RESUME

ED 410 698 EC 305 748

AUTHOR Olson, John F.; Goldstein, Arnold A.
TITLE The Inclusion of Students with Disabilities and Limited English Proficient Students in Large-Scale Assessments: A Summary of Recent Progress. Research and Development Report.
INSTITUTION National Center for Education Statistics (ED), Washington, DC.
REPORT NO NCES-97-482
ISBN ISBN-0-16-049132-0
PUB DATE 1997-07-00
NOTE 118p.
AVAILABLE FROM U.S. Government Printing Office, Superintendent of Documents, Mail Stop: SSOP, Washington, DC 20402-9328.
PUB TYPE Reports - Descriptive (141)
EDRS PRICE MF01/PC05 Plus Postage.
DESCRIPTORS *Disabilities; *Educational Assessment; *Educational Trends; Elementary Secondary Education; Evaluation Methods; *Inclusive Schools; *Limited English Speaking; Research and Development; Student Evaluation; *Testing Problems; Trend Analysis
IDENTIFIERS *National Assessment of Educational Progress; Testing Accommodations (Disabilities)

ABSTRACT

This report presents a compendium of approaches to the inclusion of students with disabilities or of Limited English Proficiency (LEP) in large-scale assessments such as the National Assessment of Educational Progress (NAEP) or state assessment programs. Chapter 1 presents an overview and lists organizations involved in this effort. Chapters 2 and 3 summarize inclusion activities and research projects related to increasing the participation of students with disabilities or LEP, respectively, in assessments. Information on the different types of accommodations used around the nation is included. Chapter 4 presents details of the approaches and procedures implemented by NAEP in recent years, focusing on the implementation of an NAEP-like assessment in Puerto Rico in 1994, the national field test in 1995 of revised procedures, and the implementation of new inclusion and accommodation procedures in the 1996 NAEP. In Chapter 5, a listing of the status of ongoing research projects studying varying aspects of the use of new procedures for NAEP is included, with a discussion of yet-to-be-resolved technical issues and future directions for NAEP and large-scale assessments in general. Appendices list acronyms, define key terms, and identify further resources. (Contains 112 references.) (DB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 410 698

NATIONAL CENTER FOR EDUCATION STATISTICS

Research and Development Report July 1997

The Inclusion of Students With Disabilities and Limited English Proficient Students in Large-Scale Assessments

A Summary of Recent Progress

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

U.S. Department of Education
Office of Educational Research and Improvement NCES 97-482

BEST COPY AVAILABLE

EC 305 748

ED 410 698

NATIONAL CENTER FOR EDUCATION STATISTICS

Research and Development Report July 1997

The Inclusion of Students With Disabilities and Limited English Proficient Students in Large-Scale Assessments

A Summary of Recent Progress

John F. Olson
Education Statistics Services Institute

Arnold A. Goldstein
National Center for Education Statistics

U.S. Department of Education
Office of Educational Research and Improvement NCES 97-482

EC 305 748

U.S. Department of Education
Richard W. Riley
Secretary

Office of Educational Research and Improvement
Ramon C. Cortines
Acting Assistant Secretary

National Center for Education Statistics
Pascal D. Forgione, Jr.
Commissioner

The National Center for Education Statistics (NCES) is the primary federal entity for collecting, analyzing, and reporting data related to education in the United States and other nations. It fulfills a congressional mandate to collect, collate, analyze, and report full and complete statistics on the condition of education in the United States; conduct and publish reports and specialized analyses of the meaning and significance of such statistics; assist state and local education agencies in improving their statistical systems; and review and report on education activities in foreign countries.

NCES activities are designed to address high priority education data needs; provide consistent, reliable, complete, and accurate indicators of education status and trends; and report timely, useful, and high quality data to the U.S. Department of Education, the Congress, the states, other education policymakers, practitioners, data users, and the general public.

We strive to make our products available in a variety of formats and in language that is appropriate to a variety of audiences. You, as our customer, are the best judge of our success in communicating information effectively. If you have any comments or suggestions about this or any other NCES product or report, we would like to hear from you. Please direct your comments to:

National Center for Education Statistics
Office of Educational Research and Improvement
U.S. Department of Education
555 New Jersey Avenue NW
Washington, DC 20208–5574

July 1997

The NCES World Wide Web Home Page address is
http://www.ed.gov/NCES/

Suggested Citation
U.S. Department of Education. National Center for Education Statistics. *The Inclusion of Students With Disabilities and Limited English Proficient Students in Large-Scale Assessments: A Summary of Recent Progress*, NCES 97-482, by John F. Olson and Arnold A. Goldstein. Washington, DC: 1997.

Contact:
Arnold A. Goldstein
(202) 219–1741

For sale by the U.S. Government Printing Office
Superintendent of Documents, Mail Stop: SSOP, Washington, DC 20402-9328
ISBN 0-16-049132-0

FOREWORD

This report is an important step in the progress of the National Center for Education Statistics (NCES) toward the study and development of an inclusive approach to large-scale assessments. The study of this topic is receiving increased attention because of a number of reasons: recent educational trends and legislation, desire for fairness and equity in assessment approaches, enhancement of samples that are more representative of the U.S. population of students and that will provide an improved measure of achievement for all students and permit better generalizability of results. In response to recommendations from advisors and experts in the field toward increasing the participation of students with disabilities and limited English proficient students in the National Assessment of Educational Progress (NAEP), NCES recently implemented several changes to the assessment and initiated a number of studies to closely examine the issues related to increased inclusion.

Because of the increasing importance of the topic, a special project was created to assist NCES in the monitoring, synthesis, and summarization of the many inclusion-related assessment activities underway in the field. This two-year study, conducted by the Education Statistics Services Institute (ESSI), is helping NCES to continuously interact with key constituent groups, collect and review a large amount of information on their progress and findings, share results and plans as they become available, and suggest future directions for large-scale assessments such as NAEP.

This report, the first of a two-volume series, describes many of the recent efforts at the national, state, and local levels to increase the participation of students with disabilities and limited English proficient students in large-scale assessments, including the efforts and progress made by NAEP. The report also serves as a compendium of approaches to inclusion, an introduction to current developments in increasing the participation of students with special needs, and a resource for continued examination of the topic. The second volume is planned for completion in late 1997. The next report will further document the studies currently underway and summarize their latest findings, as well as recommending areas for additional research on unresolved issues.

NCES will continue to closely examine the issues involved in developing the most appropriate and inclusive approaches for use in large-scale assessments. Because of the importance of inclusion, this topic requires much attention and our continuing efforts. We believe the outcomes of our efforts will be valuable to many in the field. Not only will they result in an improved national assessment, but they will also benefit educators and practitioners at state and local levels. NCES is committed to more inclusive assessments. This report is one of a continuing series of steps that indicate our commitment and progress toward this goal.

Gary Phillips
Associate Commissioner, Assessment Group
National Center for Education Statistics

ACKNOWLEDGMENTS

This report, *The Inclusion of Students with Disabilities and Limited English Proficient Students in Large-Scale Assessments: A Summary of Recent Progress*, was co-authored in an NCES/ESSI collaboration through a project of the Education Statistics Services Institute. The authors, John Olson of ESSI and Arnold Goldstein of NCES, worked as a team under the general direction of Sharif Shakrani, NAEP Program Director for Item Development and Operations in the Assessment Group at NCES, and Ramsay Selden, Director of ESSI.

The authors would like to thank the many individuals who provided invaluable assistance and guidance toward the conduct of this project and the production of this report. Many people, representing a number of different organizations, were involved in the project, and continue to be involved in activities related to the inclusion of students with special needs in large-scale assessments. They contributed greatly to various aspects of the project and the completion of this volume.

The project was funded by the National Center for Education Statistics, in the Office of Educational Research and Improvement of the U.S. Department of Education. The NCES staff, in particular, Gary Phillips, Sharif Shakrani, and Maureen Treacy, were closely involved in all aspects of this project and played a crucial role in the guidance of the study and development of this report. Other staff at NCES reviewed drafts of the report and made many important suggestions that improved the final product. These individuals were Susan Ahmed, Peggy Carr, James Houser, Andrew Kolstad, Edith McArthur, and Mary Rollefson. Work on this project also benefited greatly from the consistent support and guidance of Pat Forgione, Commissioner of NCES.

The project was conducted under a contract with the Education Statistics Services Institute, which provides program services and direct support to NCES. The staff at ESSI, especially its managers, Ramsay Selden, Laura Salganik, and Karol Krotki, helped to implement the project, provided oversight to its conduct, and coordinated program activities with NCES. A special thanks goes to Alan Vanneman, who reviewed successive drafts and made many suggestions that helped greatly to improve the report. Other ESSI staff who provided valuable assistance in the preparation and review of the report were JoAnn Blue, Douglas Butler, Rosario Castruita, and Sara Davison.

The work of this ESSI project was supported by staff at the American Institutes for Research. In addition to AIR's overall support of the ESSI program, certain staff and groups within AIR provided critical assistance. In particular, Ruth Childs helped immensely in the initial phases of the study. AIR staff also provided valuable assistance in editorial and word processing support.

Many other people, representing key organizations involved in the topic of inclusion of special needs students in assessments, provided additional input and helped to shape this volume. They are Louis Danielson and David Malouf of the Office of Special Education and Rehabilitation Services, Delia Pompa and Milagros Lanauze of the Office of Bilingual Education and Minority Languages Affairs, Judith Leonard of the Office of the General Counsel, Rebecca Fitch and Eileen Hanrahan of the Office for Civil Rights, Edward Fuentes and Gilbert Garcia of the National Institute on the Education of At-Risk Students, Larry Feinberg of the National Assessment Governing Board, John Mazzeo of Educational Testing Service, Nancy Caldwell of Westat, Inc., Charlene Rivera of the George Washington University Center for Equity and Excellence in Education, and James Ysseldyke of the National Center on Educational Outcomes.

All of these individuals contributed their wisdom and experience toward the improvement of the report. Their high professional standards, discerning eyes, and commitment to quality were crucial to the utility and relevance of this volume. The authors are indebted to the many individuals who helped on this project and/or in the creation of the report.

TABLE OF CONTENTS

FOREWORD iii

ACKNOWLEDGMENTS v

CHAPTER 1. Introduction **1**

A. Overview 1

B. Organizations Involved in the NAEP Inclusion Process 3

C. Organization of this Report 4

CHAPTER 2. Including Students With Disabilities in Large-Scale Assessments: A Summary of Projects and Activities **7**

A. Overview 7
- Issues and Concerns 8
- Types of Disabilities and Percentages of Students with Disabilities 8
- Variations in Guidelines for Inclusion of Students with Disabilities in Large-Scale Assessments 14

B. Accommodations and Adaptations Used in Large-Scale Assessments 16
- Types of Accommodations 17
- State's Use of Testing Accommodations in the Assessment of Students with Disabilities 18
- Reporting of Results for Students with Disabilities 20

C. Summary of Current and Ongoing Research Activities and Projects 20
- NCES Inclusion-Related Projects 21
- The National Academy of Education's Studies 23
- The National Center on Educational Outcomes 24
- Projects Sponsored by the Office of Educational Research and Improvement 25
- Projects Sponsored by the Office of Special Education and Rehabilitation Services 27
- The CCSSO SCASS Assessment Consortia 29
- CRESST Studies Focusing on the Assessment of Students with Disabilities 30
- Studies Conducted by the Educational Testing Service 31
- Studies by the National Academy of Sciences/National Research Council 32
- NAEP Validity Studies Expert Panel 33
- Joint Committee on Testing Practices Project on Assessing Individuals with Disabilities 33

D. Summary 34

CHAPTER 3. Including Students with Limited English Proficiency in Large-Scale Assessments: A Summary of Projects and Activities 35

A. Overview 35
Terminology 36
General Approach, Issues, and Concerns for the Assessment of LEP Students 37
Background Information on the LEP Student Population 38
Guidelines for Inclusion of LEP Students in Large-Scale Assessments 41

B. Participation of LEP Students and Use of Accommodations and Adaptations in Large-Scale Assessments 42
Participation of LEP Students in State Assessment Programs 42
State's Use of Accommodations and Adaptations in the Assessment of LEP Students 43
Reporting of Results for LEP Students 44

C. Summary of Current and Ongoing Research Activities and Projects 44
NCES Projects and Activities 45
The Office of Bilingual Education and Minority Languages Affairs 47
The National Academy of Education's Studies of Assessing LEP Students 48
The Prospects Study 49
Research of the Stanford Working Group 50
Activities and Plans of the SCASS Assessing LEP Students Projects 51
CRESST Studies of Language Issues 52
Research Activities of the GWU Center for Equity and Excellence in Education 54
Report from the National Research Council 56

D. Summary 57

CHAPTER 4. Approaches and Procedures Implemented by NAEP in Recent Years to Increase the Participation of Students with Disabilities and Limited English Proficient Students 59

A. 1995 NAEP Field Test 60
NAEP's Approach Prior to 1995 60
Degree of Exclusion of Students Identified as IEP or LEP 60
Exclusion/Inclusion Criteria used in NAEP 61
Field Test of Revised Inclusion Criteria and New Accommodations Procedures 62
Findings from the NAEP Field Test 64

B. The Puerto Rico Special Assessment 67
Factors Leading to a Special NAEP-like Assessment in Puerto Rico 67
Design and Operational Issues for the Puerto Rico Assessment of Educational Progress 67
Findings from the Puerto Rico Assessment 68

C. 1996 NAEP Assessment 69
Special Sample Design and Implementation of New Inclusion Criteria and Accommodations 70
Reporting 1996 Results 72

D. Summary 73

CHAPTER 5. Ongoing Studies, Research Needs, and Future Directions 75

A. NCES's Ongoing Research Studies on Inclusion Issues 75
Scaling Issues 76
Reporting Issues 76
Construct Validity Issues 77
Language Complexity Issues 78

B. Further Technical and Measurement Issues 78
General Measurement Questions and Technical Issues for Large-Scale Assessments 78
Appropriateness of Inclusion Criteria and Procedures 79
Impact of Accommodations on Assessment Results 80
Reporting of Disaggregated Data 80

C. Overview of the Next Volume of this Report 81

D. Conclusion 81

REFERENCES 83

Bibliography 83

APPENDICES 93

Appendix A: List of Acronyms for Key Organizations, Groups, and Programs 94

Appendix B: Definitions of Key Terms on the Inclusion of Students with Disabilities and Limited English Proficient Students in Assessments 97

Appendix C: Resources for More Information on Inclusion and Assessment Issues 101

Chapter

1

Introduction

A. Overview

B. Organizations involved in the NAEP inclusion process

C. Organization of this report

A. Overview

The purpose of this report is to describe recent efforts to increase the participation of ***students with disabilities*** and of ***limited English proficient (LEP)*** students in large-scale assessments. Inclusive assessment systems are receiving increased attention at the federal, state, and local levels, with much activity occurring in recent years. These include the efforts being made by many state assessment programs, as well as those being made by the National Assessment of Educational Progress (NAEP). Gathered together in this report is a summary of the activities of many organizations, inside and outside the Federal government, devoted to promoting greater inclusion of students with disabilities and LEP students in large-scale assessments.

The report serves as a compendium of approaches to inclusion and describes criteria for including students with disabilities and LEP students, modifications to assessment conditions and to the assessments themselves (accommodations and adaptations) that make it possible for students with special needs to participate, and ongoing research. This volume may be viewed as an introduction to current developments in inclusion in large-scale assessment, and as a resource for further exploration of the topic. References to organizations active in the area of inclusion are provided within the text, and the report concludes with extensive lists of references and resources.

This is the first volume of a series of reports. The second volume, planned to be published by the National Center for Education Statistics (NCES) in late 1997, will contain information on the results of the studies currently underway that are examining NAEP data. The second volume will also provide recommendations for further research on the issues involved. More details on the plans for Volume II are presented in Chapter 5 of this report.

NAEP's main purpose is to provide key indicators of what the nation's students know and can do. This concept means that the NAEP results should represent *all* students in the nation. This is especially important because NAEP uses a sampling approach in which all students may be included, although, in actuality, some students with disabilities and LEP students do not participate in the assessment. Recent educational trends, reflected in the Goals 2000: Educate America Act and Improving America's Schools Act (IASA), and in proposals for the reauthorization of the Individuals with Disabilities Education Act (IDEA), have called for assessments that are meaningful, challenging, and appropriate for *all* students. This call has led NCES to look closely at the procedures for assessing larger numbers of students with disabilities and LEP students.

One of the goals of this report is to describe the activities underway in NAEP toward increasing the numbers of students with disabilities or LEP students who participate in the assessment. In this report, information is provided on the changes made to the inclusion criteria, the types of accommodations now being offered, new procedures implemented in 1996, and ongoing research studies on inclusion issues.

However, as NCES staff implement new procedures for NAEP, they are also paying close attention to issues of accuracy and representativeness of the data that are collected. The process of implementing change requires a balancing of possible benefits and consequences to the assessment. Although there is a desire for inclusiveness, the consequences of a more inclusive approach in terms of its effects on validity or the personal toll on individual test takers must be evaluated carefully. In addition, as the psychometric characteristics of a more inclusive assessment are examined closely, NCES will need to balance the challenges of maintaining a valid and reliable assessment with the general goal of inclusiveness.

NCES is committed to increasing the representation of students with disabilities and limited English proficient students in its data bases. For NAEP this means, to the extent possible, maximizing participation of students with disabilities and LEP students in assessments. The NAEP sample has traditionally been drawn from lists that include all students who are in graded schools not specifically created to serve the needs of students with special needs (such as schools for individuals who are blind). Many students with disabilities and LEP students who were selected into the NAEP sample, however, were excluded from actually participating in the assessments.

There are a variety of reasons for their exclusion, such as the requirements of the Individualized Education Programs (IEP) of students with disabilities. The IEP is a legal document that reflects the decisions made by a committee concerning a student's performance level and the corresponding goals and objectives that address the areas of need. IEPs can dictate that students not be assessed at all, that they participate in only certain types of assessments, or that they only participate in assessments if certain accommodations are made. Another reason for the exclusion of a student from an assessment can be based on the judgments of school staff concerning the inability of limited English proficient students to read and understand the assessment and to be able to participate meaningfully in it.

A reason that reinforced the importance of improving the inclusion of LEP students in the assessment was highlighted with the advent of the NAEP Trial State Assessment in 1990 and continuing state-by-state assessments conducted in 1992 and 1994. Large differences in the inclusion rates of LEP students across the participating states were observed in the data. In the NAEP state-by-state comparisons, some states with high LEP populations had inclusion rates that were very different than other states with smaller LEP populations. It was likely that different rules for including LEP students in the assessment were being used, with some states using more stringent rules than others. As a result, the state-by-state comparisons may have been differentially affected and the findings not altogether comparable across all states. Another concern was the issue of how representative the assessment results were when a significant proportion of the student population (such as LEP students) had been excluded from the samples of certain states.

The intent of NAEP has always been to gain the participation of students with disabilities and LEP students. In recent years efforts have been made to make the criteria for inclusion less open to judgment and more consistently applied across states. NAEP has recently begun providing the types of accommodations and adaptations to the assessments that would make it possible for more students with disabilities and LEP students to participate. Currently, NCES is examining a variety of issues involving the inclusion of these special needs populations in NAEP. The factors leading to the implementation of these innovations are discussed in Chapter 4.

Based on a number of reasons, NCES has taken steps to increase the participation of as many of these students as possible in the assessment. Among the many benefits of this inclusive approach are improved measurement of overall student achievement, enhancement of the representativeness and generalizability of NAEP results, and greater fairness and equity. Among the challenges for NAEP are upholding its high degree of validity, maintaining its reliability, and preserving the ability to analyze and report trends in the face of changes made to procedures and in the sampled population of respondents.

B. Organizations Involved in the NAEP Inclusion Process

The effort to increase the participation of students with disabilities and LEP students is a complex undertaking. NCES has benefited from a thorough consultation process. NCES has interacted with other interested offices within the Department of Education, and with contractors and research organizations. These groups have contributed their experience, knowledge, and perspectives, and NAEP has benefited from their diverse contributions.

The offices spearheading the Department of Education's thrust to maximize inclusion of students with disabilities and LEP students in the educational mainstream have been the Office of the Under Secretary (OUS), the Office of Educational Research and Improvement (OERI), the Office for Civil Rights (OCR), the Office of Bilingual Education and Minority Languages Affairs (OBEMLA), the Office of Special Education and Rehabilitative Services

(OSERS), and the Office of the General Counsel (OGC). Representatives of these offices have worked with NCES to improve the criteria by which students with disabilities and LEP students are included in NAEP, and to devise a strategy of inclusion that would satisfy the twin objectives of greater participation and statistical integrity of the assessment. This team reviewed the results of the 1995 field test, in which many of the procedures used in the 1996 NAEP were first tried, and consulted closely with the NAEP staff to devise the 1996 assessment design.

Various organizations outside the Department of Education also have contributed greatly to NCES's efforts to increase inclusion of students with disabilities and LEP students, and NAEP has benefited from the information and informed judgment all these groups provided. These include research organizations, consortia, and survey companies. Research organizations, such as the American Institutes for Research (AIR) in support of the National Academy of Education (NAE), and the National Center for Research on Evaluation, Standards, and Student Testing (CRESST), provided evaluations of NAEP and studied how well the criteria for inclusion of students with disabilities and LEP students in the assessment were working. The National Research Council (NRC) of the National Academy of Sciences (NAS) provided guidance in developing a new strategy for inclusion. The National Center on Educational Outcomes (NCEO) has conducted numerous studies on assessment of students with disabilities, and provided guidance on inclusion criteria. The National Clearinghouse for Bilingual Education and the National Association for Bilingual Education provided input on the inclusion of LEP students. The Council of Chief State School Officers (CCSSO) collected information on inclusion practices in the states and organized state collaborative committees on assessment of students with disabilities and with LEP students. And the Stanford Working Group provided initial guidance on issues dealing with LEP students.

The primary NAEP contractors also have contributed meaningfully to increasing the participation of students with disabilities and LEP students in the assessment. Under the direction of NCES, Educational Testing Service (ETS) and Westat, Inc. were charged with making operational the revised inclusion criteria, as well as both testing and then incorporating the accommodations and adaptations that have made it possible for larger numbers of students with special needs to participate. These organizations were fully involved in planning and implementing the inclusion effort.

C. Organization of This Report

This report is organized into five chapters. Chapter 1 presents an overview of the topic of including students with disabilities and LEP students in large-scale assessments, and lists many of the organizations involved in this effort. Chapter 2 provides a summary of inclusion activities and research projects related to increasing the participation of students with disabilities in assessments. Information on the different types of accommodations used around the nation is included. Chapter 3 provides a summary of inclusion activities for LEP students, with a discussion of the types of adaptations used to assess these students. Chapter 4 presents details of the approaches and procedures implemented by NAEP in recent years,

focusing on the implementation of a NAEP-like assessment in Puerto Rico in 1994, the national field test in 1995 of revised procedures, and the implementation of new inclusion and accommodation procedures in the 1996 NAEP. In Chapter 5, a listing of the status of ongoing research projects studying various aspects of the use of new procedures for NAEP is included, with a discussion of yet-to-be-resolved technical issues and future directions for NAEP and large-scale assessment in general. Several appendices are included that provide additional information.

Chapter 2

Including Students With Disabilities in Large-Scale Assessments: A Summary of Projects and Activities

A. Overview
- Issues and concerns
- Types of disabilities and percentages of students with disabilities
- Variations in guidelines for inclusion of students with disabilities in large-scale assessments

B. Accommodations and adaptations used in large-scale assessments
- Types of accommodations
- State's use of testing accommodations in the assessment of students with disabilities
- Reporting of results for students with disabilities

C. Summary of current and ongoing research activities and projects
- NCES inclusion-related projects
- The National Academy of Education's Studies
- The National Center on Educational Outcomes
- Projects sponsored by the Office of Educational Research and Improvement
- Projects sponsored by the Office of Special Education and Rehabilitation Services
- The Council of Chief State School Officers SCASS assessment consortia
- CRESST studies focusing on the assessment of students with disabilities
- Studies conducted by the Educational Testing Service
- Studies by the National Academy of Sciences/National Research Council
- NAEP Validity Studies Expert Panel
- Joint Committee on Testing Practices project on assessing individuals with disabilities

D. Summary

A. Overview

In recent years, attention has increasingly focused on a myriad of issues related to the inclusion of students with disabilities in assessments. Legal requirements, such as Section 504 of the Rehabilitation Act of 1973 and its implementing regulations prohibit discrimination against individuals with disabilities. Title III of the Americans with Disabilities Act of 1990 requires entities to provide accommodations to individuals with disabilities, regardless of whether they receive Federal funds. These laws were motivating

factors that helped to focus attention on the need for accommodations, although they did not focus specifically on inclusion in assessments. More recently, this approach is reflected in educational reform laws such as the Goals 2000: Educate America Act, the Improving America's Schools Act, and in proposals to reauthorize the Individuals with Disabilities Education Act. These acts helped to raise the level of effort required of educators and policymakers to comply with current law and provide additional assistance to students with disabilities.

Issues and Concerns

The exclusion of students with disabilities from national and state assessments is an issue of increasing concern. During the past half decade, attention has focused on the inclusiveness of assessment approaches, and the policies and practices implemented by various testing programs. One of the initial findings, as researchers began to examine more closely the issues of including students with disabilities in assessments, was that most assessment programs did not have much data on the excluded students. In many cases, students with disabilities were not included in the assessments. For example, one study found that, although some of these assessments were not entirely appropriate for a small percentage of students (approximately 2 percent of the student population), a large percentage of excluded students (approximately 85 percent of students with disabilities) were judged to be capable of taking the assessments (NCEO SR17 1995). Some of these students were probably capable of taking the assessments without any modifications, and others needed various accommodations.

This chapter presents a general overview of issues involved in the inclusion of students with disabilities in large-scale assessments, as well as information on the numbers of students affected by different types of disabilities and recent trends in the data. The main issues discussed focus on the guidelines and criteria used for the inclusion of students with disabilities and the lack of consistency across states in their use. Another issue concerns the different types of accommodations and adaptations that are available to assess students with disabilities and their use in different localities. Finally, a summary of the activities and projects underway in a variety of different organizations is presented.

Types of Disabilities and Percentages of Students With Disabilities

Individuals who are categorized as students with disabilities are a heterogeneous and diverse group. The Office of Special Education Programs (OSEP) reports data on 12 categories of students with disabilities, and students may display a wide range of abilities and skills within each category. In part, this may be a reflection of differences in policies and practices used to identify students with disabilities in different areas across the country. Based on NAEP data from 1994, approximately 5 percent of all students in grades four and eight were excluded on the basis of an Individualized Education Program that specified the extent of the student's participation in an assessment. An IEP can call for (1) the student not to be assessed, (2) the student to only participate in certain types of assessments, or (3) the student to only participate if certain accommodations are provided. Note that all students covered under IDEA have IEPs. Also, some students with disabilities are covered under Section 504, but not

IDEA. These students generally have plans that describe the kinds of services they will receive, which may include information about testing accommodations.

As described in the Eighteenth Annual Report to Congress (U.S. Department of Education, OSEP 1996), data are provided on the numbers of students with different types of disabilities for the most recent school year, as well as trend data over the past several years. Data are based on information from the states on the number of children served under the IDEA, Part B legislation. According to OSEP, the IDEA, Part B programs served 4,915,168 students ages 6–21 during school year 1994–95. The number of students in each disability category is reported in table 2.1. Overall, the number of students with disabilities ages 6–21 years increased 2.8 percent from school years 1993–94 through 1994–95. In comparison, the total number of students ages 6–21 in the U.S. population, according to Census data estimates for the school year 1994–95, was 58,315,764. This was an increase of only 0.89 percent from the total of 57,803,809 in school year 1993–94.

Increases within several of the disability categories were proportionately greater than the 2.8 percent increase overall for all students with disabilities (see table 2.1). The largest increase occurred in the traumatic brain injury category, which increased by 33.2 percent (from 5,395 to 7,188). Significant increases also occurred in the categories of other health impairments (28.2 percent from 83,080 to 106,509) and autism (19.5 percent from 19,058 to 22,780). There was a significant decrease in the category of students with multiple disabilities (from 109,730 to 89,646, or -18.3 percent). Some of these increases probably do not reflect additional numbers of students with disabilities. Rather, they are in part due to shifting classifications and expansion of service populations[1].

As further described in the OSEP report, students with specific learning disabilities continue to account for more than half of all students with disabilities (51.1 percent). During the school year 1994–95, 2,513,977 students with specific learning disabilities were served under IDEA, Part B, 3.5 percent (85,915) more than in school year 1993–94 under the Part B and Chapter 1 Handicapped Programs (see table 2.1). However, the school year 1994–95 percentage of students with learning disabilities in the resident population ages 6-21 is identical to the school year 1993-94 percentage. Students with speech or language impairments (20.8 percent), mental retardation (11.6 percent), and serious emotional disturbance (8.7 percent) made up an additional 41.1 percent of all students ages 6-21 with disabilities. Again, these percentage distributions are similar to the school year 1993–94 distributions.

[1] According to the OSEP report, the increases in the number of students with autism and traumatic brain injury are probably due to the relative newness of those reporting categories. School year 1994–95 was the third year states were required to report the number of students in those categories (reporting was optional for those categories in school year 1991–92). Many states attributed these increases to the provision of technical assistance to districts on the identification and evaluation of students with autism and traumatic brain injury. States also indicated that during triennial review and evaluations, these relatively new categories were likely used for students who previously were reported under other disability categories. The increase in the number of students with other health impairments appears to be the result of an expansion of the service population. Many states indicated that the increase was primarily due to increased service provision to students with attention deficit disorder. This is the third year several states have reported increases in the number of students identified as having other health impairments because of increased services to students with attention deficit disorder.

Table 2.1—Number and Percentage Change of Students Ages 6–21 years Served Under Part B and Chapter 1 (SOP): School Years 1993–94 through 1994–95

Disability	Total		Change		Percent of Total for 1994–95
	1993–94	1994–95	Number	Percent	
Specific learning disabilities	2,428,062	2,513,977	85,915	3.5	51.1
Speech or language impairments	1,018,208	1,023,665	5,457	0.5	20.8
Mental retardation	553,869	570,855	16,986	3.1	11.6
Serious emotional disturbance	415,071	428,168	13,097	3.2	8.7
Multiple disabilities	109,730	89,646	-20,084	-18.3	1.8
Hearing impairments	64,667	65,568	901	1.4	1.3
Orthopedic impairments	56,842	60,604	3,762	6.6	1.2
Other health impairments	83,080	106,509	23,429	28.2	2.2
Visual impairments	24,813	24,877	64	0.3	0.5
Autism	19,058	22,780	3,722	19.5	0.5
Deaf-blindness	1,367	1,331	-36	-2.6	0.0
Traumatic brain injury	5,395	7,188	1,793	33.2	0.1
All disabilities	4,780,162	4,915,168	135,006	2.8	100.0

NOTE: For school year 1993–94, funding for children and youth with disabilities included children counted under IDEA, Part B and the Chapter 1 Handicapped Program. For school year 1994–95, all children were counted under IDEA, Part B.

SOURCE: U.S. Department of Education, Office of Special Education Programs, Data Analysis System (DANS). Table slightly modified from original.

In table 2.2, data from the OSEP report are provided by state on the numbers and percentage change of students served under IDEA, Part B, for all types of disabilities. Note that, prior to October 1994, children and youth with disabilities in state-operated institutions were served under the IDEA, Part B, and Chapter 1 of ESEA. In October 1994, Congress passed the Improving America's Schools Act, which consolidated funding for children and youth with disabilities in state-operated institutions under IDEA, Part B. Data reported in this table for years prior to 1994 include children served under Chapter 1.

The data in table 2.2 show that the percentage change in the number of children with disabilities increased substantially from the school year 1993–94 to 1994–95 in a number of states. Those with the largest increases were Oregon, Minnesota, Nevada, and Tennessee. Most states had increases in the number of students served under IDEA, although some showed a decline.

Table 2.2—Number and Percentage Change of Students Ages 6–21 Years, by State

	Total Number Served		Change in Number Served	% Change in Number Served
State	1993–94	1994–95	1993–94 to 1994–95	1993–94 to 1994–95
Alabama	90,599	90,673	74	0.08
Alaska	15,373	15,484	111	0.72
Arizona	61,845	65,185	3,340	5.4
Arkansas	46,215	45,736	-479	-1.04
California	481,746	492,400	10,654	2.21
Colorado	59,358	61,398	2,040	3.44
Connecticut	63,988	66,835	2,847	4.45
Delaware	13,243	13,414	171	1.29
District of Columbia	6,394	6,289	-105	-1.64
Florida	256,893	269,431	12,538	4.88
Georgia	111,274	116,431	5,157	4.63
Hawaii	13,358	13,938	580	4.34
Idaho	19,914	19,888	-26	-0.13
Illinois	222,944	226,416	3,472	1.56
Indiana	115,087	117,511	2,424	2.11
Iowa	56,740	58,355	1,615	2.85
Kansas	44,017	45,805	1,788	4.06
Kentucky	66,871	66,750	-121	-0.18
Louisiana	75,848	79,053	3,205	4.23
Maine	26,477	27,345	868	3.28
Maryland	85,980	87,719	1,739	2.02
Massachusetts	139,112	142,403	3,291	2.37
Michigan	161,503	165,231	3,728	2.31
Minnesota	78,125	83,217	5,092	6.52
Mississippi	58,257	59,095	838	1.44
Missouri	104,900	108,851	3,951	3.77
Montana	16,270	16,044	-226	-1.39
Nebraska	33,475	34,743	1,268	3.79
Nevada	22,027	23,463	1,436	6.52
New Hampshire	20,793	21,758	965	4.64
New Jersey	171,978	175,970	3,992	2.32
New Mexico	39,843	41,248	1,405	3.53
New York	319,454	329,352	9,898	3.10
North Carolina	121,471	124,419	2,948	2.43
North Dakota	11,104	11,057	-47	-0.42
Ohio	203,528	205,447	1,919	0.94
Oklahoma	66,503	65,839	-664	-1.00
Oregon	57,353	61,296	3,943	6.87
Pennsylvania	187,110	187,721	611	0.33
Puerto Rico	39,196	37,179	-2,017	-5.15
Rhode Island	20,784	21,562	778	3.74
South Carolina	71,359	72,722	1,363	1.91
South Dakota	13,389	13,528	139	1.04
Tennessee	107,347	113,928	6,581	6.13
Texas	373,858	389,893	16,035	4.29
Utah	46,694	46,650	-44	-0.09
Vermont	9,281	9,536	255	2.75
Virginia	117,328	123,420	6,092	5.19
Washington	88,862	91,653	4,791	5.52
West Virginia	39,230	40,857	1,627	4.15
Wisconsin	86,764	89,165	2,401	2.77
Wyoming	10,569	10,655	86	0.81
Total	4,769,631	4,903,958	134,327	2.82

SOURCE: U..S. Department of Education, Office of Special Education Programs, Data Analysis System 1996

BEST COPY AVAILABLE

Also presented in the OSEP report are data on the five-year trends in the number of students served (see table 2.3). As reported, the number of students ages 6–21 years served increased by 12.7 percent (553,417) from school year 1990–91 through 1994–95. The largest increase occurred in the number of students with other health impairments, which increased by 89 percent (from 56,349 to 106,509). As noted earlier, much of the increase may be related to students with attention deficit disorder. A large increase also occurred in the category of students with orthopedic impairments (22.8 percent from 49,340 to 60,604).[2]

Specific learning disabilities increased by 17.3 percent (from 2,144,017 to 2,513,977). There were also increases in the new categories of autism (from 5,415 to school year 22,780) and traumatic brain injury (from 245 to 7,188). Two categories have decreased since 1990-91: deaf-blindness (-12.7 percent from 1,524 to 1,331) and multiple disabilities (-8.2 percent from 97,629 to 89,646).[3]

[2] Note that, for a number of years, Michigan has combined the orthopedic impairments category with the other health impairments category. Students in both of these categories are reported under students with orthopedic impairments.

[3] The decrease in the multiple disability category occurred in school year 1994–95 as a result of a decision by the Wisconsin SEA to report all students by their primary disability condition.

Table 2.3—Number of Students Ages 6–21 years Served During School Years 1990–91 through 1994–95

Disability Condition	School Year					Change from 1990–91 through 1994–95	
	1990-91	1991-92	1992-93	1993-94	1994-95	Number	Percent
Specific learning disabilities	2,144,017	2,247,004	2,366,487	2,428,112	2,513,977	369,960	17.3
Speech or language impairments	987,778	998,904	998,049	1,018,208	1,023,665	35,887	3.6
Mental retardation	551,457	553,262	532,362	553,869	570,855	19,398	3.5
Serious emotional disturbance	390,764	400,211	401,652	415,071	428,168	37,404	9.6
Multiple disabilities	97,629	98,408	103,279	109,730	89,646	-7,983	-8.2
Hearing impairments	59,211	60,727	60,616	64,667	65,568	6,357	10.7
Orthopedic impairments	49,340	51,389	52,588	56,842	60,604	11,264	22.8
Other health impairments	56,349	58,749	66,063	83,080	106,509	50,160	89.0
Visual impairments	23,682	24,083	23,544	24,813	24,877	1,195	5.0
Autism	NA	5,415	15,580	19,058	22,780	22,780	--
Deaf-blindness	1,524	1,427	1,394	1,367	1,331	-193	-12.7
Traumatic brain injury	NA	245	3,960	5,395	7,188	7,188	--
All disabilities	4,361,751	4,499,824	4,625,574	4,780,212	4,915,168	553,417	12.7

NOTES: The data for school years 1990–91 through 1993–94 include children ages 6-21 years of age served under IDEA, Part B and Chapter 1 Handicapped Program. For school year 1994–95 all children ages 6–21 years are served under Part B, which includes children previously counted under the Chapter 1 Handicapped Program. Autism and traumatic brain injury were introduced as separate reporting categories in school year 1991-92 as a result of P.L. 101-476, the 1990 Amendments to IDEA.

SOURCE: U.S. Department of Education, Office of Special Education Programs, Data Analysis System (DANS).

The NCES report, *The Condition of Education 1996* (U.S. Department of Education 1996), presented recent data on the education of students with disabilities. As described in the report, schools are providing many more disabled students, particularly those with learning disabilities, with special services. The IDEA mandates that all children have available to them a free and appropriate public education designed to meet their unique needs. Changes in the number and distribution of students with disabilities affect the level of effort required of educators and policymakers to comply with the current law and help them to forecast the need for future resources. Following are some findings related to these issues from *The Condition of Education 1996* report.

- The number of students participating in federal programs for children with disabilities has been increasing at a faster rate than total public school enrollment. Between 1977 and 1994, the number of students who participated in federal programs for children with disabilities increased 46 percent, while total public school enrollment decreased 2 percent.

- The percentage of disabled students identified as having specific learning disabilities rose 24 percentage points (from 22 to 46 percent) between 1977 and 1994, while the proportion identified as mentally retarded or with speech or language impairments each fell 16 percentage points (from 26 to 10 percent and from 35 to 19 percent of the total, respectively).

- The ratio of the number of students with specific learning disabilities per special education teacher serving them increased from 18 to 24 students per teacher between 1977 and 1993. However, the ratio for all students with disabilities decreased over the same period; in 1977, there were 21 students per teacher, and in 1993, 16 students per teacher.

Variations in Guidelines for Inclusion of Students with Disabilities in Large-Scale Assessments

Prior to 1990, administrations of NAEP, as well as the National Education Longitudinal Study of 1988 (NELS:88), relied on the judgment of school administrators as to whether or not the student could take the assessment. However, these criteria were not considered to be rigorous enough for determining which students with disabilities should take the assessment (Houser 1995).

Beginning with the 1990 NAEP, schools were given guidelines informing them that they may exclude a student with a disability if "the student is mainstreamed less than 50 percent of the time and is judged incapable of participating meaningfully in the assessment, OR, the IEP team or equivalent group had determined that the student is incapable of participating meaningfully in the assessment." Schools were instructed to include students with

disabilities if school staff believed the students were capable of taking the assessment. Schools were also instructed that when there was doubt, students should be included.[4]

Although written guidelines on the participation of students with disabilities in statewide assessments exist in most states, they vary a greatly from one state to another and little consistency exists across states in their requirements. Based on the NCEO's experiences working with states, the guidelines used by states may be based on any of the following criteria (NCEO SR17 1995):[5]

- IEP function—In the majority of states, the role of the IEP team is key in making the decision to include the student. In most states, the IEP document must include a statement about the participation of the student in the statewide assessment.

- Role of parents—Some states' guidelines refer directly to the role of parents in decisions on participation. States with high stakes assessments, such as exit exams, have requirements that the parent/guardian must be informed that nonparticipation in the assessment means the students will not receive a regular diploma.

- Acceptability of partial testing—Several states address the possibility of partial testing, (i.e., taking part of the test, but not all of it, such as the mathematics portion but not the reading portion) in their guidelines. This approach has been recommended as a way to help increase the participation of students with disabilities in educational accountability systems (Ysseldyke & Thurlow, 1994).

- Extent to which decisions are based on category of disability or placement—Some states use guidelines that recommend decisions about participation be based on the category or placement of the students with the disability.

- Assessing what is taught—Several states caution against the measurement of what the student has not had the opportunity to learn, and emphasize the importance of assessing student achievement only if the student has been taught the topics being assessed and at their instructional level.

- High stakes assessments vs. other assessments—A number of states require students to pass an assessment in order to receive a high school diploma. State guidelines refer to a disabled student's right to attain a regular diploma, and participation decisions in these high stakes assessments are made by the student's instructional planning team.

---

[4] Note that, for the 1996 NAEP, these criteria were changed to be more inclusionary. The criteria referring to mainstreaming was dropped and the new focus was primarily on participation as specified in the IEP and the permitted use of accommodations. Further details on the new inclusion criteria are provided in Chapter 4.

[5] Note that most of the states are in the process of reviewing and changing their guidelines for participation and inclusion. Specific details of each state's guidelines are presented in the document *A Compilation of States' Guidelines for Including Students with Disabilities in Assessments* (NCEO SR17 1995).

- Alternate means of assessment—A few states have made available different forms of assessment for students whom the regular statewide assessments are inappropriate. The intent is to select an assessment that will give a true measure of the student's proficiency, especially in cases where the standardized procedures are viewed as not appropriate for some students.

- Reporting of results—Many states describe what they will do with data on students with disabilities. These data include results of standard administrations of statewide assessments, accommodated administrations of assessments, and use of alternate assessments. States that report data have varying guidelines about what data are reported, and whether the data are aggregated in reports.

B. Accommodations and Adaptations used in Large-Scale Assessments

Research on the use of testing accommodations involves looking at a dynamic and continuously changing field. Much work on the topic has been done over the past few years, and the rate of change in the use of accommodations has been dramatic. In the 1993 NCEO synthesis report, *Testing Accommodations for Students with Disabilities: A Review of the Literature*, the authors concluded that not enough was known about acceptable accommodations and no set of appropriate guidelines existed (NCEO SR4 1993). Their report summarized the uses made of various testing accommodations, adaptations, and modifications. The literature review was organized into four topic areas: (1) policy and legal considerations, (2) technical concerns, (3) minimum competency and certification/licensure testing efforts, and (4) existing standards. This report, one of the earliest to examine the issues related to the use of accommodations, compared practices across states and found many inconsistencies. Because of the newness of the topic at that time, many more questions than answers existed, leading to a call for answers to the most pressing questions. Most immediately, the report recommended that a comprehensive set of guidelines be developed that state and national agencies could use in decision making. These guidelines addressed issues of (1) inclusion/exclusion criteria, (2) how and when to modify tests or testing procedures, and (3) how to report scores and summarize data.

The terms "accommodation," "adaptation," and "modification" are often used interchangeably, however, some uses conveying different meanings (NCEO SR4 1993). In the assessment literature, testing "modifications" or "adaptations" are usually associated with changes made specifically to the test format. Some examples are Braille booklets, large-print versions, audiocassette tapes, and the use of readers for students with visual impairments. Assessment "accommodations" are usually associated with changes in the testing environment. Examples include taking the assessment in a different setting, such as individually instead of in a group, or under flexible time arrangements, such as allowing the student unlimited time to take the test. Many practitioners in the field use the terms interchangeably, and no real consensus exists on specific uses of the terms. In this report, the terms accommodations and adaptations are primarily used, as they are defined above.

The provision of accommodations is closely connected to decisions to include or exclude students with disabilities from large-scale assessments. In order to increase the number of students with disabilities who are included, the use of reasonable accommodations in the assessments must be considered (NCEO SR18 1995). In the early 1990s, concerns about the lack of adequate accommodations for students with disabilities were expressed, and many assessment programs increasingly began to implement them. At the national level, for the first time in 1995, NAEP conducted field tests to examine the possible use of different types of accommodations in future administrations of the assessment. Some were subsequently used in the operational assessment in 1996.

Recently, states have been making much progress in developing guidelines for the use of accommodations in assessments. Increasing numbers of states are making provisions for students with disabilities to use accommodations as part of their statewide assessments. According to a recent state survey conducted by the NCEO, the number of states who have guidelines for the use of accommodations had increased from 21 in 1992 to 39 in 1995 (NCEO SR18 1995). The survey found that the use of accommodations in statewide assessments was quite prevalent, but not very consistent. States were found to vary greatly in the detail of their accommodations guidelines and in what they permitted. Most states have revised their guidelines for the use of accommodations in recent years.

Types of Accommodations

A variety of accommodations have been identified and are being used in large-scale assessments. Assessment accommodations can be grouped together into categories depending on how they are used. For example, the NCEO has categorized some of the most common assessment accommodations, as listed below.

Accommodations related to timing, including:

- Allowing for extended time
- Providing more breaks during testing
- Arranging for extended testing sessions over several days

Accommodations in the assessment environment that enable the student to:

- Take the test alone, in a testing carrel
- Take the test in a small group
- Take the test at home
- Take the test in a special education class

Modifications to the response format that allow the student to:

- Mark responses in the assessment book and not on answer sheet
- Point to the response
- Give oral responses
- Give responses in sign language
- Use template for responding

- Use typewriter for responding
- Use computer for responding
- Get personal assistance in making responses, either oral or written

Modifications to the presentation format that include the use of:

- Reading the directions out loud
- Interpretation of directions
- Use of magnifying equipment
- Signing of directions
- Braille editions
- Large-print versions

These accommodations can be categorized in one of two types of classifications—those that are modifications to the administration process, such as changes in the assessment environment, timing, or response format, and those that are modifications to the presentation format or the assessment instrument itself, such as Braille editions or other language versions. Some types of accommodations often require that multiple accommodations be allowed, such as reading aloud of directions, which requires additional time to administer.

States' Use of Testing Accommodations in the Assessment of Students with Disabilities

The CCSSO and the North Central Regional Educational Laboratory (NCREL) conduct an annual survey of state assessment programs and practices. Data from the Association of State Assessment Programs (ASAP) survey, conducted in 1995, were used to determine how many students with disabilities currently participate in statewide assessment programs and what kinds of special testing conditions or accommodations were allowed to enable these students to participate. The findings that follow are discussed in the article *Statewide Assessment of Students with Disabilities* (Bond 1996).

Data from the 1996 State Student Assessment Programs (SSAP) database were also made available, the fourth year that these data were collected and reported. Information is available on the numbers of states operating assessment programs of different types, the purposes and types of measures used in the programs, and the subject areas assessed. Data were also collected on emerging trends and new areas of increased interest, such as the use of standards to develop and report assessment information and plans for the use of statewide assessment for Title I purposes. Data from the report *The Status Report of Assessment Programs in the United States*: *State Student Assessment Programs Database* (Bond et al. 1996) are included below.

As of 1995, most states (45) had implemented some type of statewide assessment program; only four states had no program. A majority of the states (41) had written guidelines about the participation of students with disabilities in their programs. Although not all states were able to estimate participation rates for students with disabilities, for those that could, the rates for students with disabilities ranged from 6 to14 percent of the total tested elementary school student population and from 5 to10 percent of the total tested secondary school

population. State test directors and special education directors question the accuracy of these numbers because the data are either not collected in many places, or they are not always gathered in a systematic way that provides an accurate estimate of the participation rates of special education students (Bond 1996).

Information was also gathered in the ASAP survey on the policies states use to determine if students with disabilities participate in the state assessment programs. In most states, a special education student is included or excluded from the state assessment based on the recommendations made in the student's IEP. But in a few states, the decision is based on criteria related to the student's reading grade level. A few states (such as California, Idaho, Michigan, Utah) use a 50 percent rule (i.e., include the student in the assessment if the student spends 50 percent or more of his or her time in regular education classes in the tested subject), although in these states the IEP may be used to override this rule.

Many states allow for special testing conditions and accommodations. Although most states are very willing to allow for accommodations that can enable students with physical or sensory disabilities to take the test, for example Braille versions for blind students, problems may arise for disabilities that are cognitive or affective in nature. Some accommodations for cognitive problems provide students with extra help that may or may not be related to the subject being tested. For example, although it may be valid to read the questions on a mathematics test to a student who has dyslexia or some other reading disability, it may not be valid to have someone read a test of reading to the same student. The problem arises when the accommodation is closely related to the skill being assessed. In this case, the student's mathematics score would be a more accurate estimate of his or her mathematics knowledge than the reading score would be of his or her reading ability (Phillips 1994, 1995).

From the SSAP data, 37 states reported the use of special testing accommodations. Table 2.4 reports the testing accommodations states allow for students with disabilities. All, or almost all, permitted the use of Braille and large-print versions of the test, small group administrations, flexible scheduling, extra time, and separate test administrations. Some states, such as Maryland and Hawaii, provide numerous accommodations, including Braille and large print versions, reading or transcribing the test, extended time periods, small group administration, audiotaped versions, signed versions for the hearing impaired, use of calculators, and use of word processors. A number of states mentioned that decisions concerning special accommodations depended on their impact on the validity and interpretability of the results. It should be noted that, although states have developed lists of accommodations they are permitted in testing situations, data on the numbers of students who actually use these accommodations are not readily available. The data in table 2.4 focus on what states permit and not on what is actually used by students in the states. It is probable that there are differences between what is permitted and what is used.

Table 2.4—Permissible Accommodations for Students with Disabilities

Type of Accommodation	Number of States that Permit Accommodation
Large Print	34
Braille or Sign Language	33
Small Group Administration	33
Flexible Scheduling	31
Separate Testing Session	31
Extra Time	30
Audiotaped Instructions/Questions	27
Multiple/Extra Testing Sessions	25
Word Processor	21
Simplification of Directions	15
Audiotaped Responses	12
Other Accommodation	12
Use of Dictionaries	9
Alternative Test	6
Other Languages	2

SOURCE: The Status of State Student Assessment Programs in the United States (CCSSO/NCREL 1996)

Reporting of Results for Students with Disabilities

Another issue related to participation of students with disabilities in statewide assessment programs is the reporting of their scores. This issue is especially critical for accountability systems where the intent is to be accountable for all students. Many states offer schools the option of not reporting, or reporting separately, the scores of the special education students in the state, district, and school averages. Although in these localities, the policy may be to include as many special education students as possible, the concern is that the results may not be totally comparable to those of other students because of the special circumstances used to assess these students. Even though an exact number was not available from the SSAP database, many states indicated that the assessment results of students with disabilities may be eliminated from state, district, and school assessment summaries.

C. Summary of Current and Ongoing Research Activities and Projects

Increasing amounts of information about the assessment of students with disabilities are available as more research activities investigate the topic. Still, many questions need to be answered as progress is made in studying further the issues and modifying the approaches used to increase the participation of students in assessments. Some of these questions are:

- Are as many students with disabilities being included in large-scale assessments as possible?

- Are adequate and appropriate accommodations provided?
- Is the content of the assessments appropriate for students with particular types of disabilities?
- Do students with disabilities study the same curricula as other students?
- Are the assessments valid for these students?
- What impact do accommodations, such as allowing extended time to take the assessment, have on the scores for students with disabilities? What impact do they have for non-disabled students?

Currently, these and other questions are being examined, with additional studies being planned or just getting underway. Many different people and organizations are involved in efforts to understand better the issues surrounding attempts to increase the participation of students with disabilities in large-scale assessments. NCES is working closely with many of these groups, following the ongoing activities and research underway on the topics of inclusion and assessment accommodations. NCES hopes to benefit from the findings of these projects, and use them to improve their own surveys and assessments. In this section, summaries of the main projects and activities are provided.[6]

NCES Inclusion-Related Projects

NCES is currently conducting and coordinating a number of studies about the participation of students with disabilities in NAEP. Several of the studies are being conducted under the operational contract for NAEP, and additional studies are being conducted by other groups researching various aspects of including students with disabilities and LEP students in the assessment. More details of these NAEP-specific studies are presented in Chapters 4 and 5.

Working Conference on Guidelines for Inclusion of Students with Disabilities and Accommodations in Large-Scale Assessment Programs. In March 1994, NCES held a working conference in Washington to discuss a wide variety of issues related to the inclusion and accommodation of students with disabilities in national and large-scale assessments. In addition, guidelines were developed for making decisions on including and accommodating students, and technical and implementation issues were identified that form a research agenda on the issues. Staff from NCES and OSERS attended the conference, along with staff from other organizations involved in the issues.

Following the conference, the report *Making Decisions About the Inclusion of Students with Disabilities in Large-Scale Assessments: A Report on a Working Conference to Develop Guidelines on Inclusion and Accommodations* (NCEO SR13 1994) was disseminated. The

[6] The authors of this report acknowledge that some important activities may have been excluded in this summary, and apologize to those who may have been inadvertently omitted. The intent here is to list the key activities related to NCES' programs, and in particular, NAEP.

report summarized the discussions and decisions made at the NCES meeting, as well as recommending guidelines for immediate implementation and future directions. The main recommendations were for NCES to develop more objective criteria for schools to use when making decisions about inclusion and exclusion, and for NCES to initiate an active research agenda to investigate data validity issues related to including students with disabilities and LEP students in their assessments. Included in the report were also recommendations to:

- Include students with disabilities when trying out items to be included in an assessment.
- Include all students with disabilities in some form of the assessment. This includes partial participation in the assessment or use of an alternative assessment for some students.
- Include students with disabilities in the reporting of results.
- Although not all students with disabilities need accommodations during assessments, they should be used when needed to increase the number of students with disabilities who can take the tests.
- As new technologies and procedures for accommodations and adaptations are developed, they should be subjected to validation research.
- Adherence to the intent of the guidelines needs to be monitored. Strong incentives are needed so no student is excluded who could participate if accommodations and adaptations were used.

Many of the decisions made at the meeting were incorporated into NAEP's procedures for the upcoming assessments. For example, the exclusion criteria were revised, accommodations were provided, and additional students were able to participate in 1996. Details are presented in Chapter 4.

Working Paper on Assessing Students with Disabilities and Limited English Proficiency. In early 1995, NCES prepared a working paper, *Assessing Students with Disabilities and Limited English Proficiency* (Houser 1995), which discussed the current state of affairs as it related to NCES policies that resulted in excluding some students from assessments and surveys. Highlighted were concerns about the validity of the data NCES collects if a portion of the population of students were excluded, the potential of bias in the results of surveys that did not include students with disabilities or limited English proficiency, and the effects of using accommodations or alternative assessments. Additional points made in the report were as follows:

- Many students with disabilities who were currently being excluded from NCES assessments could be included with no accommodations.

- For students requiring accommodations, care needs to be taken not to stigmatize any student who would require special accommodations.

- In this document it was speculated that two possible categories of accommodations exist—ones that do not affect test scores and ones that may affect scores. Issues related to the effect of accommodations on scores from large-scale assessments need to be studied carefully.

- Test scores based on different types of accommodations may need to be reported separately if it is determined they are not psychometrically equivalent to one another (i.e., when various accommodations affect the results differently).

- As modifications of the approaches used by NCES are made, care must be taken not to invalidate NCES's measurement of trends.

The working paper concluded with a description of current studies that are being conducted by NCES on including students with disabilities and LEP students in NAEP. Because these studies were just getting underway as part of the 1995 NAEP field test, no results were available in the working paper. Results from these studies are presented in Chapter 4.

The National Academy of Education's Studies

Since 1990, several studies have been carried out by the American Institutes for Research under a contract with NAE to evaluate the NAEP Trial State Assessment (NAE 1993, 1996). In the NAE evaluations of the 1992 and 1994 assessments, a focus was on issues related to inclusion of students and accommodations in the assessment. In the 1993 report, a study of students excluded from the trial state assessment was conducted (NAE 1993). The study found that more students identified as IEP could probably have participated in the assessment, especially if accommodations or alternative assessments had been available. The panel recommended that additional students with IEPs should be included in future NAEP assessments, but suggested that special studies be done first as part of the 1994 assessment to determine the severity of disabilities among students who had been excluded from NAEP and to make a judgment of the assessability of these students. In addition, it recommended that NCES should examine the cost effectiveness of shifting resources away from getting large sample sizes in order to reduce sampling error and move towards including students who are more expensive to assess, (e.g., students with disabilities and LEP students), in order to reduce sampling bias.

Findings based on the 1994 Trial State Assessment of reading suggest that 83 percent of the fourth grade students with IEPs would have been assessable on the current NAEP reading instrument on the basis of their reading scores, including 70 percent of the students who had been excluded in 1994. However, the authors concluded, many of the students with disabilities read below their current grade level, and there are not enough items on the NAEP

reading assessment to assess accurately those students who perform at the lower end of the proficiency distribution.

The 1994 studies examined further the assessability of students with IEPs that were excluded. Teachers were asked about the kinds of accommodations that would be needed to allow excluded students to participate (Stancavage et al. 1996). These studies also examined the assessability of LEP students, which is discussed further in Chapter 3. The NAE Panel found that teachers appeared to be using the appropriate criteria correctly for their decisions to exclude students (i.e., based on the student's reading level and the percent of time mainstreamed), although they tended to assume that participation required a higher level of reading proficiency than was found in this study. In addition, it was found that teachers were very likely to recommend assessing students with disabilities under accommodated conditions. Even though no accommodations were available for the 1994 assessment, from the teachers' perspective, accommodations would have been beneficial for more than half of the students with disabilities, including almost two-thirds of those who had been included in the 1994 assessment, as well as over 40 percent of those who had been excluded. The most frequently recommended accommodations were for extended time and/or shorter tests. Teachers also recommended oral reading of directions and testing in small groups for a large percentage of the students. The authors advised that this recommendation would then lead to an increase in the total number of students with disabilities assessed, but a substantial decrease in the number assessed under standard conditions. This would occur because some students who might take the assessment without accommodations would use accommodations if offered.

The National Center on Educational Outcomes

NCEO was established in 1990 to provide leadership in the identification of outcomes, indicators, and assessments to monitor educational results for all students, including students with disabilities. In particular, the NCEO is involved in working with states and federal agencies to identify important outcomes for students with disabilities, examining the participation and use of accommodations by students with disabilities in national and state assessments, evaluating national and state practices in reporting assessment information on students with disabilities, and studying the availability and use of statewide assessment results for students with disabilities.

In many ways, the NCEO can be considered the leader in the collection of data and the publication of information on the topic of inclusion of students with disabilities and accommodations for assessments. They have produced an extensive list of publications that include technical reports, state activity updates, and policy documents (see references for a list of relevant NCEO publications). Information from some of their reports has been discussed earlier in this chapter. In particular, the following NCEO reports are especially useful in providing valuable summaries and detailed information on the topic: *Recommendations for Making Decisions About the Participation of Students with Disabilities in Statewide Assessment Programs* (NCEO SR15 1994), *Compilation of State Guidelines for Including Students with Disabilities in Assessments* (NCEO SR17 1995),

Compilation of State Guidelines for Accommodations in Assessments for Students with Disabilities (NCEO SR18 1995), and *Assessment Guidelines that Maximize the Participation of Students with Disabilities in Large-Scale Assessments: Characteristics and Considerations* (NCEO SR25 1996).

Included in this collection of reports are recommendations for including and accommodating students with disabilities in state assessments, and for the reporting of their results. Also, compilations of guidelines summarizing states' approaches and general policies for students of special populations, applicable laws, participation of students in assessment programs, the types of tests used by the states are listed, as well as descriptions of the written laws, regulations, and guidelines that states have about the use of accommodations in statewide assessments. In addition, criteria for use in evaluating state or district assessment guidelines for students with disabilities are provided.

Currently, NCEO is working with the CCSSO and the Association of State Testing Directors, with funding from OSERS, to identify topics that need further study and to develop a program of collaborative studies and full-scale development projects that focus on inclusion activities and the use of accommodations. In this capacity, NCEO has compiled a directory of assessment projects and related efforts on the inclusion of students with disabilities (NCEO 1996). This directory summarizes current activities being conducted by a number of different organizations, but primarily includes those projects sponsored by OERI or OSERS. In the directory, NCEO identified many major research efforts or activities at the state and national levels focusing on different aspects of the assessment of students with disabilities. Eight projects, funded by OERI, are focusing on activities in state assessments. These involve projects in seven states, plus one project being done as a multi-state consortium with the CCSSO. Three projects, funded by OSERS, focus on the participation and accommodation of students with disabilities in performance and alternative assessments. Summaries of plans for the projects, with an emphasis on the assessment of students with disabilities, are included in following two sections.

Projects Sponsored by The Office of Educational Research and Improvement

OERI offers assessment development grants under the National Institute for Student Achievement, Curriculum, and Assessment. In 1996, a number of state projects were funded, which include work on creating, at the state level, better assessments for students with disabilities, examining the issues related to providing appropriate accommodations, and adapting or modifying statewide assessments for use with either students with disabilities or LEP students. Details of these projects are given below.

- Inclusive Comprehensive Assessment System Project. This project of the Delaware Department of Public Instruction involves the development of an inclusive assessment system in language arts, mathematics, social studies, and science. The system will be designed to best meet the needs of students with disabilities and LEP students in mathematics at grades 3 and 8, and science at grades 5 and 10. The initial focus will be on identifying effective assessments for students with learning

disabilities. In ensuing years, effective assessments will be identified for other, lower incidence, disability groups. The project also will include the development of Spanish-embedded tasks and linguistic accommodations for on-demand tasks to measure the academic progress of Spanish special education LEP students and those whose first language is other than Spanish or English.

- Maryland Assessment System Project. The Maryland State Department of Education is working on a project that includes an examination of the validity and reliability of the state's School Performance Assessment System, with a focus on issues related to the inclusion and performance of special education students. Research will examine the impact of test administration accommodations for special education students on score validity for the state assessment and the relationship between classroom and assessment accommodations. Currently, several studies are being conducted on accommodations in high stakes performance assessment. Issues that are being investigated include the effects of commonly used accommodations on the reliability and validity of the test, the relationship between classroom and assessment accommodations, and the identification of other accommodations that could be used but are not currently available. In addition, the first report from this project, *A Comparison of State Assessment Systems in Maryland and Kentucky with a Focus on Participation of Students with Disabilities*, is forthcoming.

- Modifying Minnesota Graduation Standards for Students Who do not Speak English and for Students with Disabilities. The Minnesota Department of Education is conducting this project, whose overall goal is to develop, implement, and evaluate policies and procedures to ensure that students with disabilities and LEP students can fully participate in the major reform effort underway in the state's high school competency test. The project will design approaches to ensure that all students participate in the new assessments and develop policies and a decision-making process to determine participation and appropriate accommodations for students with disabilities and LEP students. As the state prepares for the first administration of its new exams, focus groups of teachers and administrators are exploring the best ways to include and accommodate students with disabilities. Staff members have produced annotated bibliographies related to the assessment of students with disabilities and LEP students, and the information is currently under review.

- North Dakota Language Arts Assessment. This project, being conducted by the North Dakota Department of Public Instruction will develop valid, reliable, and generalizable assessments of the language arts that can provide parents, students, teachers, administrators, and the community with accurate and easy-to-understand indicators of each student's progress in reaching content standard benchmarks and performance standards. The intent is to develop assessments that are equitable and fair for all students, including students with disabilities and LEP students.

- Oregon Assessment Development and Evaluation Project. The Oregon Department of Education is working on a project that involves the development of assessments in

science, and will evaluate the current statewide assessments in reading, writing, and mathematics, with a focus on adaptations and modifications for students with disabilities and LEP students to ensure that all students participate. Spanish versions of the assessments will be field tested and evaluated for use with LEP students. Adaptations of the assessments will be field tested and evaluated for use with students with disabilities. The project will also conduct research on the validity and reliability of Oregon's state assessments, including comparability and fairness of the assessments. Analyses were being conducted on the effects of two accommodations—(1) a study of the differences for fourth graders between marking in the test bubbles vs. filling in the bubbles on an answer sheet, and (2) a study of the differences for eighth graders between response made by hand vs. by word processing.

- Pennsylvania Assessment through Themes Project. The Pennsylvania Department of Education is involved in a project that will design, develop, and prepare for the implementation of performance assessment measures in the content areas of science and the arts. The goal is to develop a statewide longitudinal assessment plan that will progress toward an integrated, cross-discipline assessment after the year 2000. Comprehensive, theme-based performance assessments will be produced by exploring methods of assessing all students in science and the arts through the use of themes. The project will also identify methods for modifying the standards and assessments for students whose primary language is not English, and for students with special needs or with disabilities.

- State Collaborative on Assessment and Student Standards (SCASS) Technical Guidelines for Performance Assessments (TGPA) Standards. This project involves a consortium of about 20 states, with the CCSSO coordinating the work. The group is working collaboratively to address the issues of performance assessment and the inclusion of students with disabilities and LEP students. The project focus includes the evaluation of assessments for all students, including special education and LEP students. The project includes methods and plans for disseminating information resulting from project activities to all 50 states and other interested organizations. Practical guidelines will be created for the design, development, and implementation of alternative assessment procedures. More details on the SCASS TGPA are provided in the following section.

Projects Sponsored by the Office of Special Education and Rehabilitation Services

The Division of Innovation and Development (DID) of the Office of Special Education Programs (OSEP) is sponsoring some 250 projects involving research on the education of individuals with disabilities. These projects are funded by the DID and authorized under the IDEA. Currently, three of the projects are investigating issues related to the assessment of students with disabilities. As described in *Innovation and Development in Special Education: Directory of Current Projects* (U.S. Department of Education OSEP/DID 1996), these projects are:

- Examining Alternatives for Outcome Assessment for Children with Disabilities. The Maryland State Department of Education is involved in another project (in addition to the one listed in the previous section) to conduct studies on assessment accommodations and alternative assessment approaches that facilitate the participation of students with disabilities. Plans are for 12 studies that will form a systematic research program on accommodations and alternative assessments. Included in these projects are an exploration of how assessment systems must be developed to have the range and flexibility to accommodate all students, including those with disabilities; studies of the relationships between instructional accommodations and accommodations in assessments; issues and concerns of reliability and validity in the use of testing accommodations and adaptations, alternative assessments, and specification of standards and outcomes; and the use of alternative assessments for students with severe cognitive disabilities.

- Performance Assessment and Standardized Testing for Students with Disabilities Project: Psychometric Issues, Accommodation Procedures, and Outcome Analyses. The Center for Education Research at the University of Wisconsin—Madison is conducting a project whose purpose is to study how students with disabilities react to and perform on on-demand and in-class performance assessment tasks in science and mathematics. Specifically, the project will focus on the use of performance assessment tasks from the statewide student assessment system of mathematics and science, and how testing accommodations procedures function for students with disabilities. Data on psychometric issues, accommodations procedures, and the utility of different types of assessments for evaluating academic outcomes will be examined, as well as information from interviews conducted with students and questionnaire data gathered from teachers. Following the 1996 assessment in Wisconsin, students and teachers were interviewed about the assessment, and the accommodations used by each student with disabilities were identified and rated by teachers as to their fairness and helpfulness.

- Project Reading ABC— An Alternative Reading Assessment Battery for Children with Severe Speech and Physical Impairments. The purpose of this project, being done by the Center for Literacy and Disability Studies at Duke University, is to develop an alternative reading assessment battery for children with severe speech and physical disabilities. An interdisciplinary team of experts in assessment, literacy, and severe impairments will systematically design and validate a reading instrument for use with these children. Performance of children without disabilities and those with severe disabilities will be compared on alternative and more traditional measures of the reading construct.

The CCSSO State Collaboratives on Assessment and Student Standards Assessment Consortia

Begun by the CCSSO in 1991, the SCASS program was created to provide a mechanism for states to work together to develop assessments related to standards. Consortia of states were created to explore ways of developing assessment designs in specific subject areas. States join to participate in the various SCASS projects and work together to share their expertise and resources in order to develop improved assessments in less time and at reduced costs. At this time, over 40 states are involved in one or more of the on-going SCASS development projects of assessment consortia. Of particular interest to those involved in the topic of inclusion of students with disabilities and LEP students in assessments are three SCASS groups. These are consortia on assessing special education students, technical guidelines for performance assessment (both discussed further below), and assessing LEP students (discussed further in Chapter 3).

SCASS Consortium on Assessing Special Education Students. In early 1996, a group of states formed a consortium to focus on the inclusion of special education students in large-scale assessment programs. This collaborative is co-directed by the CCSSO, NCEO, and the National Association of State Directors of Special Education. The purposes of this new consortium are for states to share existing methods and criteria for accommodating special education students in large-scale assessment programs, to determine the state of the art regarding both accommodation and adaptation of assessments, and to plan a research-based program to develop criteria and procedures so that the performance of all students on state standards can be assessed. Products and services of the consortium are to assist states and local districts to develop a comprehensive set of materials and procedures for adaptation of assessments, to design procedures that allow the maximum number of special education students to participate in the regular large-scale assessment programs (either with or without accommodations), and to identify data collection procedures for use with the small numbers of students who may still be unable to participate at all. Materials and guidelines for their use will be developed at the state level, and these will then be adapted for use at the local level. In addition, monitoring and compliance procedures will be instigated at the state and local levels to assure that the materials and procedures are being used appropriately.

The Special Education SCASS group met several times in 1996, and began making progress in a number of critical areas. The group identified 19 key issues that need further study. Study groups were formed to address the following three issues: (1) participation rules, incentives, and actions that increase the participation of students with disabilities in assessments; (2) development of training packages for teachers and parents on making participation and accommodation decisions; and (3) development of a paper on the implications of including scores from students with disabilities in aggregated results versus reporting the scores separately.

SCASS Consortium on Technical Guidelines for Performance Assessment. The SCASS Technical Guidelines for Performance Assessment (TGPA) consortium is focused on designing and implementing practical research to foster the development of technically

BEST COPY AVAILABLE

sound performance assessments. Research studies sponsored by the consortium use state data and resources to address issues common to participating states. In 1996, this SCASS group received a grant from the U.S. Department of Education to support the design and implementation of research, the development of practical guidelines for developing and implementing large-scale performance assessment, the compilation and dissemination of research on performance assessment, and professional development seminars and materials for state assessment staffs.

In 1996, research was underway on four projects, with plans for addressing five additional topics in various stages of development and review. The group is developing guidelines for performance assessment and is collecting reports of research on performance assessment, encouraging the use of such assessments in large-scale programs. Research relevant to the issues of inclusion of students with disabilities includes the SCASS TGPA "Study 6—The Impact of Adapting Assessments for Special Education Students." The goal of this project is to study the impact of adapting large-scale assessments for special education students, and to give states guidance on which adaptations are likely to be most appropriate.

The SCASS TGPA met several times in 1996, and work has begun on designing a multi-state study on the effects of assessment accommodations on student performance. The current focus is on the effects of reading the assessment to the students on tests in subjects other than reading, and response accommodations involving oral responses and technical assistance. A pilot test is scheduled for the spring of 1997.

CRESST Studies Focusing on the Assessment of Students with Disabilities

The National Center for Research on Evaluation, Standards, and Student Testing (CRESST) initiated two studies in 1996 to investigate issues on including students with disabilities and LEP students in assessments. In this section, a brief summary of plans for the study on students with disabilities are presented; plans for the LEP study are included in Chapter 3. As presented by Dan Koretz at the September 1996 CRESST conference, the study of students with disabilities will focus primarily on validity issues and the quality of measures of disabled students' performance. Among the issues to be examined are characteristics of the assessment population, the difficulty level of current assessments that are being used to assess these students, what types of accommodations seem most reasonable, and the validity of the results from accommodated assessments.

According to Koretz, the project will include descriptive studies of the characteristics and performance of students with disabilities, how these students are currently being assessed, and the effects of including students with disabilities in statewide assessments. Data on the percentages of students identified as disabled will be compared across states. One concern is that many states are using assessments that are targeted more for high achieving students and may not be as appropriate for those students with disabilities who are performing at the lower end of the achievement scale. A critical issue is the design of the accommodations, with the goal being to offset any possible bias or distortions caused by the student's disability. Possible difficulties in designing accommodations involve the use of ambiguous

BEST COPY AVAILABLE

disability categories leading to an unclear determination of the students with disabilities, the interaction of effects when some students have more than one type of disability, interference effects due to the accommodation of disabilities that are related to the test constructs being measured, and a serious lack of research findings on the topic. Preliminary results from this study will be available in 1997.

Studies Conducted by the Educational Testing Service

ETS has been one of the primary organizations involved in conducting in-depth research on measurement issues related to the inclusion of students with disabilities in tests. Since the 1980s, the ETS Committee on People with Disabilities has helped to develop policies on the testing of students with disabilities, coordinated studies of the issues, and disseminated a number of documents related to the topic of accommodating and including students with disabilities (see references for additional ETS reports on this topic). These issues have been a part of policy and practice for many standardized testing programs ETS (e.g., SAT, GRE) for some time.

The report, *Testing Persons with Disabilities: A Report for ETS Programs and their Constituents* (ETS undated), summarized the findings of a four-year study on the effectiveness of various test accommodations and their impact on the meaning of the resulting scores. Accommodations offered included alternative test versions, assistive personnel, assistive devices, separate testing locations, and extra time. A large section of the report dealt with the issue of extra time, which is an important issue for standardized tests that are speeded. Unfortunately, no definitive findings were provided on exactly how much extra time should be allowed and how additional time may affect the validity of the test scores. Nevertheless, ETS encouraged the provision of ancillary test materials in alternate formats and special scoring services comparable to those available to nondisabled examinees.

During the 1980s, ETS researchers conducted numerous studies focusing on measurement and validity issues and the testing of disabled students (Bennett et al. 1985, 1987, 1988, 1989; Braun et al. 1986; Ragosta et al. 1986; Rock et al. 1987, 1988; Willingham 1987, 1989). The book *Testing Handicapped People* (Willingham et al. 1988) contained results of many of these studies, which examined a variety of issues on the topic, for example, characteristics of the population, performance on standard and nonstandard tests, use of accommodations, admissions decisions, psychometric characteristics, and effect on validity. In research on accommodations for the SAT and GRE, it was indicated that standard and accommodated versions of the tests were generally comparable in terms of reliability, factor structure, differential item functioning, prediction of academic performance, admissions decisions, and test content. However, providing additional time for students with learning disabilities may present problems with data validity, especially for certain types of tests where time is a factor. It was found that extra time adversely affected the predictive validity of the test, in that college academic performance tended to be overpredicted for these students (Willingham et al. 1988). Students with learning disabilities who received additional time to take the SAT or GRE did not perform as well academically as their test

BEST COPY AVAILABLE

scores predicted they would. Test scores for students with learning disabilities who received additional time biased the data and overpredicted their postsecondary grades by approximately one-third of a standard deviation. The authors questioned the policy of allowing extra testing time to complete the test for all students with learning disabilities, regardless of the severity or nature of the disabilities.

In 1995, ETS sponsored a seminar on "The ADA and its Impact on Standardized Testing," in which the issues brought about by the ADA and their impact on providers of standardized tests were examined. The meeting focused on the provisions of the ADA related to standardized testing and review of recent court decisions, issues related to the administrations of standardized tests to individuals with disabilities, and new approaches that are being taken to address equity of access. ETS is considering the possibility of holding another seminar in the near future to discuss disability testing issues for policymakers and researchers.

Studies by the National Academy of Sciences/National Research Council

The NAS is currently involved in two studies related to the inclusion of students with disabilities in nationwide assessments and in state reform plans. The first study, which began in 1995, was mandated by Congress in the Goals 2000: Educate America Act. Margaret McLaughlin and Lorraine McDonnell are serving as co-chairs. The Committee on Goals 2000 and the Inclusion of Students with Disabilities is examining a wide variety of issues on the topic. These issues include the social and legislative background; implications for practice; the impacts and costs of reform; and implications for research and policy. Evidence that is being examined by the committee includes available databases and other data sources, various reports and other published information, how Goals 2000 is being implemented in states' plans, and relevant legal and policy research. A report draft was to be finalized by the end of 1996 and a final version will be disseminated in June 1997.

More recently, the NAS received a 5-year grant to conduct an evaluation of NAEP. The NAS Committee on the Evaluation of National and State Assessments of Educational Progress was convened in 1996 by Congressional mandate, with the authority to conduct the evaluation granted by Title IV, National Education Statistics, Section 411. The committee has been charged with reviewing NAEP generally and evaluating the developmental state assessments, student performance levels, and the extent to which results are reasonable, valid, and informative to the public. Of particular interest, is the committee's examination of sampling and data collection methods, including the inclusion of second-language learners and students with special learning needs, such as students with disabilities. According to the committee's plans, special consideration will be given to the interpretation of data for these examinees and the meaningfulness of the results.

The committee plans to sponsor a workshop in 1997 to further the thinking of policymakers, sponsors, and other NAEP users. With the oversight of the NRC's Board on Testing and Assessment, the committees involved in the evaluation of NAEP will issue recommendations and an evaluation report to the Secretary in the fall of 1998.

BEST COPY AVAILABLE

NAEP Validity Studies (NVS) Expert Panel

The NVS Panel is in the process of planning a study on validity issues related to providing accommodations to students with disabilities. The focus of this proposed study will be on the effects of extra time. As discussed in the draft "A Proposed Study Design to Examine the Impact of Accommodations of the Performance of Students with Disabilities in NAEP: The Impact of Increased Testing Time on the Performance of Disabled and Non-Disabled Students" (Olson 1997) to the NVS Panel, the impact of testing accommodations on student performance needs to be better understood in order to know whether the scores of accommodated students can be combined in a valid way with those of other students. Questions related to the impact of different types of accommodations on student performance need to be answered in order to determine what accommodations are needed for students with disabilities to help them participate in an assessment without unduly impacting test validity.

The proposal suggests a study to examine the impact of extended testing time on students with and without disabilities, the impact of extended testing time across different content areas (e.g., reading and mathematics), and student perceptions of their need for accommodations. The intent of this study is to assess the impact of a particular accommodation (extended time) on students with and without disabilities. The study is tentatively planned to be conducted in 1997, possibly beginning with a preliminary study of existing NAEP data from the 1996 assessment or the 1997 field test. However, since this is a proposal under review, plans and timelines are subject to change.

Joint Committee on Testing Practices (JCTP) Project on Assessing Individuals with Disabilities

The JCTP is an inter-professional association coordinating group that is involved in a project to improve the quality of testing practices. Participating associations include the American Counseling Association, American Psychological Association, American Speech-Language-Hearing Association, National Association of School Psychologists, and the National Council on Measurement in Education. Various testing organizations are a part of the collaboration through the participation of their staff in JCTP projects and disseminators of their products. Past products include the "Code of Fair Testing in Education," "Test User Qualifications: A Data-Based Approach to Promoting Good Test Use," "Responsible Test Use: Case Studies for Assessing Human Behavior," and the "ABCs of School Testing".

A working group of the JCTP is currently working on a project that will compile information about what is being done regarding assessment of individuals with disabilities. Their goal is to develop a sourcebook that provides information helpful to test users, such as assessment specialists, counselors, educators, personnel specialists, psychologists, and providers of services related to special education. A special focus will be on information for policymakers to use when they have questions about the assessment of individuals with disabilities, about interpreting test scores from such assessments, and about making

educational and career decisions using information from these high-stakes assessments. The JCTP's project on Assessing Individuals with Disabilities plans to have the sourcebook available in early 1998.

D. Summary

This chapter focused on issues concerning the inclusion of students with disabilities in large-scale assessments. Background data were provided on the numbers and types of students with disabilities. According to data from the Office of Special Education and Rehabilitation Services, during the school year 1994–95, there were about five million students aged 6 to 21 with disabilities served under Part B of the IDEA. Students with learning disabilities, speech or language impairments, mental retardation, or a serious emotional disturbance make up the bulk of this group.

Recent studies have found that many students with disabilities who were excluded from large-scale assessments at the federal and state levels were, in fact, capable of participating. Some of these students could only have participated in assessments if accommodations had been available. Many large-scale assessments, such as NAEP and some state assessments, are now offering accommodations to students who need them. The accommodations may vary, but often include use of large-print booklets, Braille or sign language, testing in multiple sessions or in small-group or individual sessions, flexible scheduling, and extended time. However, at this time, there is no nationally accepted set of guidelines for provision of accommodations.

Various criteria for deciding whether students with disabilities can participate in large-scale assessments are in use around the country. Although written guidelines exist in most states, they vary among locations. Many assessment administrators rely on the student's Individualized Education Plan for guidance, but others may take additional factors into account, such as parental input, student placement, and availability of alternate assessments.

Several research projects are underway studying the topic of inclusion of students with disabilities. The main issues being studied include the following: the degree to which these students are excluded from testing; use of accommodations; appropriateness of assessment content for students with disabilities and alignment of assessment content with what special education students are studying; inclusion of special education students in state assessment programs; and validity of assessments administered under nonstandard, or accommodated, conditions.

In the next chapter, a summary of the issues and activities concerning the inclusion of limited English proficient students in large-scale assessments is presented.

Chapter 3

Including Students with Limited English Proficiency in Large-Scale Assessments: A Summary of Projects and Activities

A. Overview
- Terminology
- General approach, issues, and concerns for the assessment of LEP students
- Background information on the LEP student population
- Guidelines for inclusion of LEP students in large-scale assessments

B. Participation of LEP students and use of accommodations and adaptations in large-scale assessments
- Participation of LEP students in state assessment programs
- State's use of accommodations and adaptations in the assessment of LEP students
- Reporting of results for LEP students

C. Summary of current and ongoing research activities and projects
- NCES projects and activities
- The Office of Bilingual Education and Minority Languages Affairs
- The National Academy of Education's Studies of Assessing LEP Students
- The Prospects Study
- Research of the Stanford Working Group
- Activities and plans of the SCASS Assessing LEP Students Project
- CRESST studies of language issues
- Research activities of the GWU Center for Equity and Excellence in Education
- Report from the National Research Council

D. Summary

A. Overview

The education of students with limited English proficiency have been a focus of policymakers, educators, and the courts for the past three decades. According to federal law, as well as many state laws, if students cannot participate meaningfully and equitably in an English-only school environment because of their LEP status, they are eligible for special services. The types of programs set up to serve the needs of these students can vary tremendously. The range of instructional programs for LEP students include some that

initially encompass substantial instruction in the native language, with gradual transition to English over a course of several years. A small number of programs seek to develop proficiency of students in both English and a second language. Some students may receive a proportion of their instruction exclusively in English that has been simplified, along with an instructional content that has been enriched to make the content more understandable. Other students may receive their instruction in English as a second language (ESL), with a primary focus on the development of English language skills rather than on the academic content areas.

The main issues connected with the assessment of LEP students concern data availability, participation, and validity of assessment results. Data on educational outcomes can be difficult to obtain for LEP students since many are usually not being tested. A critical concern is the effect of language on the validity of achievement measures administered in English, if the measures are administered at all. For example, in the past, NAEP had systematically excluded many students who were LEP and were judged incapable of taking the assessment in English. In another NCES survey, the National Educational Longitudinal Study of 1988, only 327 students identified as LEP were included in the sample, but 3,831 LEP students were eliminated because it was judged that they were insufficiently proficient in English to take the test or complete the questionnaires.

Because of this, concerns exist that any estimates based on the sample of English-language learners who participate in the assessment or took the test may be more representative of those most proficient in English. These concerns are compounded because testing those most proficient in English will result in only a small percentage of all LEP students being tested. In addition, once a student is proficient in English, he or she is no longer considered LEP, so the percentage of LEP students most proficient in English tends to be small.

Terminology

Many different terms are in use for the types of students and the programs designed to serve them. According to OBEMLA, students who come from language backgrounds other that English are "language minority students." References to language minority students pertain to individuals from homes where a language other than English is frequently used, and therefore have an opportunity to develop a level of proficiency in a language other than English. Students who are language minority students and whose proficiency in English has not yet developed to the point where they can fully participate in an English-only instructional environment are "limited English proficient (LEP)." The term LEP is used in most national and state data collection efforts, federal and state education legislation, as well as in court cases involving the rights of these students. According to some researchers in the field, the term "English language learner (ELL)" is viewed as being a more positive label than is LEP, and is preferred in their writings (August and Hakuta 1997; GWU/CEEE 1996; LaCelle-Peterson and Rivera 1994).

In addition to these terms, another one is also used for certain purposes. Programs that are designed to serve the needs of language minority students and use the students' native

language as they acquire the necessary skills in English often use the term "bilingual." In bilingual education programs, the proportion of native language used often varies greatly from program to program. Some bilingual education programs have as a goal bilingual proficiency, but not all do. For example, the model known as the transitional bilingual education program uses the native language only to teach English as quickly as possible, and not with the goal of maintaining or improving native language proficiency. Bilingual students can be defined as individuals who are proficient in a native language other than English who have developed enough proficiency in English not to be disadvantaged in an English-only school environment (August and Hakuta 1997).

In this report, the term LEP is used primarily, although other terms like language minority students or bilingual may be used when referring to these students or to special types of programs.

General Approach, Issues, and Concerns for the Assessment of LEP Students

In general, assessment approaches for limited English proficient and bilingual children have included: (1) identification of students whose English proficiency is limited and (2) determination of eligibility for participation in the assessment program. Once a student is determined to be eligible, decisions must be made as to the type of approach to use to assess the student. Of particular concern to national, state, and local assessments is how the validity and reliability of assessments apply to LEP and bilingual children and the appropriateness of the assessments for these children. Use of any assessment for LEP students must meet standards of validity and reliability, as stated for professional practice (American Educational Research Association, American Psychological Association, and the National Council on Measurement in Education 1985).

States and school districts use a variety of methods to determine which language minority students are limited English proficient and to monitor the progress of the students in special programs in which they may be placed. These methods include registration and enrollment records, home language surveys, interviews, observations, referrals, classroom grades and performance, and test results (Cheung et al. 1994). The administration of English language proficiency tests is the most common method. Hopstock and Bucaro (1993) found that 83 percent of local districts used English language proficiency testing, either alone or in combination with other techniques, to determine which language minority students were limited in their English proficiency. In general, there is a great amount of variability across school districts in the way assessments are used to identify and place LEP students. Many states provide overall guidance to districts on assessment procedures for LEP students but allow considerable flexibility in their choice of assessment methods and instruments. A practice increasingly being recommended is the use of a combination of assessments to obtain multiple sources of information on a variety of criteria related to the monitoring and placement of LEP students.

Background Information on the LEP Student Population

According to the 1990 U.S. Census, approximately 6.3 million school aged children (ages 5–17 years) lived in a home where a language other than English was spoken. This is about 14 percent of the total number of students in the U.S. population. A subset of this number were LEP students, but there is some disagreement on the exact number. Based on the self-reported responses from a sample of households in the census, it was estimated that approximately 900,000 children spoke English "not well or not at all," 1.5 million spoke it "well," and 3.9 million spoke it "very well."

A more direct estimate of the number of LEP students in the nation was based on a nationally representative sample of school districts done in 1991, which found the number in grades K-12 to be slightly more than 2.3 million (Fleischman and Hopstock 1993). This estimate represented an increase of almost 1 million over the results of a similar survey done in 1984. Other estimates of the size of the LEP student population have ranged from 2 million to 3.3 million (Hopstock and Bucaro 1993).

The Office of Bilingual Education and Minority Languages Affairs annually summarizes information submitted by State Education Agencies in the *Survey of States' Limited English Proficient Students and Available Educational Programs and Services* (NCBE, 1996). The 1994-1995 summary report collected information on the number of LEP students in the various states and outlying territories and jurisdictions and the educational services provided or available to them. Among the main findings presented for 1994-95 were:

- The number of LEP students enrolled in public and nonpublic schools continued to increase. The states reported 3,184,696 LEP students in 1994-95, an increase of 4.8 percent over 1993-94.

- The total reported number of LEP public school students comprised 7.3 percent of the reported public school enrollment of students in grades K-12, and 1.1 percent of the non-public K-12 enrollment.

- There is no single, nationally consistent definition for limited English proficiency across the states. Most states reported a definition of LEP that was based on a combination of non-English language background and/or difficulties with speaking, reading, writing, and understanding English.

- Most states used tests and other assessment methods to identify LEP students. Almost all states indicated that they used more than one method for this identification.

Table 3.1 shows the distribution of LEP students by grade level in public schools. This table is reproduced from the Descriptive Study report (Fleischman and Hopstock 1993). As seen in the table, larger proportions of LEP students are found more in the lower elementary

grades, with slightly more than half found in grades K-4. Decreasing proportions of LEP students are found as the grades get higher.

Table 3.1—Number and Percentage of LEP Students by Grade Level

Grade Level	Number of LEP Students	Percentage of LEP Students in Grade Level	Total Number of Students in U.S.	Percentage of LEP Students in Total
Kindergarten	277,914	12.1	3,305,619	8.4
1st grade	279,257	12.1	3,554,274	7.9
2nd	246,979	10.7	3,333,285	7.4
3rd	221,936	9.6	3,333,285	6.7
4th	197,211	8.6	3,312,443	6.0
5th	177,412	7.7	3,268,381	5.4
6th	150,421	6.5	3,238,095	4.6
7th	134,907	5.9	3,180,120	4.2
8th	125,849	5.5	3,019,826	4.2
9th	159,208	6.9	3,310,290	4.8
10th	137,101	5.9	2,913,951	4.7
11th	103,337	4.5	2,642,554	3.9
12th	75,423	3.3	2,390,329	3.2
Ungraded	16,469	0.7	---	---
Total	2,303,425	100.0	42,000,343	5.5

NOTE: Data based on district mail survey. The number of respondents was 735; this was 98.7% of those who responded to the survey. The results are weighted to be nationally representative.
SOURCE: Fleischman and Hopstock 1993

LEP students tend to be concentrated in a small number of states that have large populations (see table 3.2). Based on data from the 1990 census, 67 percent of the language minority population lived in five states—California, Texas, New York, Florida, and Illinois.

Table 3.2—Top Five States with Largest Percentage of Students Who Spoke a Language Other than English in the Home and Were Rated as Speaking English Less than Very Well

State	Number of Students Who Spoke a Language Other than English in the Home and were Rated as Speaking English Less than Very Well	Percentage of Students Who Spoke a Language Other than English in the Home and were Rated as Speaking English Less than Very Well
California	796,905	30
Texas	391,881	15
New York	247,948	11
Florida	113,441	6
Illinois	102,031	5

NOTE: Census data based on the number of children ages 5–17 years in 1990 who spoke a language other than English in the home and who were rated as speaking English less than "very well." Total number of students in this age range was 2,388,243.
SOURCE: U.S. Bureau of the Census 1990.

According to the data, as presented in table 3.3, the largest proportion of LEP students are native speakers of Spanish, who make up about three-fourths of the population. Other languages with a sizable proportion of students are Vietnamese, Hmong, Cantonese, Cambodian Korean, Laotian, Navajo, Tagalog, and Russian. Percentages of other language groups are each less than 1 percent.

Table 3.3—Number and Percentage of LEP Students in Most Common Language Groups

Language Group	Number of LEP Students	Percentage of LEP Students
Spanish	1,682,560	72.9
Vietnamese	90,922	3.9
Hmong	42,305	1.8
Cantonese	38,693	1.7
Cambodian	37,742	1.6
Korean	36,568	1.6
Laotian	29,838	1.3
Navajo	28,913	1.3
Tagalog	24,516	1.1
Russian	21,903	0.9

NOTE: Data based on district mail survey. The number of respondents was 733; this was 98.4% of those who responded to the survey. The results are weighted to be nationally representative.
SOURCE: Fleischman and Hopstock 1993

Guidelines for Inclusion of LEP Students in Large-Scale Assessments

There are many similarities between factors that lead to the exclusion of LEP students and those that result in the exclusion of students with disabilities. As described in the NCEO report described in Chapter 2, *Making Decisions about the Inclusion of Students with Disabilities in Large-Scale Assessments* (NCEO SR13 1994), many of the same issues that are posed for the inclusion of students with disabilities in large-scale assessment are the same for LEP students. Some states use similar guidelines for including or excluding LEP students as specified by NAEP. In general, the types of guidelines in effect vary from one state to another, with little consistency shown across the states.

The NAEP procedures used prior to 1990 allowed schools to exclude sampled students if they were LEP and if local school personnel judged the students incapable of participating meaningfully in the assessment. Beginning in 1990, NAEP instructions to schools for excluding LEP students from the assessment required the following conditions to be met: the student is a native speaker of a language other than English AND the student has been enrolled in an English-speaking school for less than two years (not including bilingual education programs) AND school officials judged the student to be incapable of taking the assessment. The guidelines also stated that when in doubt, the student was to be included in the assessment. These criteria were changed beginning with the 1995 field test of NAEP, discussed further in Chapter 4.

In the 1992 NAEP, approximately four percent of the fourth–grade students and three percent of eighth–grade students were identified as LEP, and about two-thirds to three-fourths of them were excluded from the assessment. In other words, three percent of all fourth graders and two percent of all eighth graders were excluded because of language barriers. In addition, the guidelines also contributed to differential exclusion rates across states participating in the NAEP Trial State Assessment, which may have had an effect on state-by-state comparisons (Spencer 1994). This probably happened because the guidelines were open to somewhat subjective interpretations by local staff who may have taken a degree of latitude in their following of the criteria as it was specified by NAEP.

Other research has identified several factors, including the exclusion guidelines used, that may have an impact on the exclusion of LEP students from large-scale assessments. These include the following (August et al. 1994):

- A lack of clear and consistent definitions of LEP at the national or state levels.
- Guidelines that exclude students who have been in bilingual education programs, even when they have been in English-speaking schools for more than two years.
- The varying degrees of English proficiency that students in bilingual programs have.
- Guidelines that allow local decisions to be made about the participation of LEP students.

- The differential implementation of guidelines.
- The failure to monitor the extent to which the intent of the guidelines is followed.
- The lack of accommodations in assessment materials and procedures that would enable some LEP students to participate.
- A desire not to require LEP students to take an assessment they cannot understand because of limited English proficiency.

In their paper, August et al. (1994) discussed the implications of increasing inclusion of LEP students in assessments such as NAEP, and the concerns that come into play when modifications are made to existing exclusion guidelines. Guidelines for developing a coherent framework for inclusion based on a number of principles related to the proper conduct of a large-scale assessment are presented. These principles include use of a continuum of strategies and assessment options, use of realistic and practical options, equity issues, and maximizing inclusion of LEP students in the assessment.

Another important area, sometimes overlooked when examining factors that may have an impact on the exclusion of LEP students, is the issue of including these students in the design and piloting of new assessments. A "best practice" approach in the development of assessment instruments is to include samples of all types of students who will be included in the assessment. Unfortunately, practices in some national and state assessments have involved including LEP students in the assessment, but not always in the piloting or field testing. This may have contributed to concerns regarding validity issues. In the future, attempts will be made to include LEP students in the field testing of new NAEP items.

B. Participation of LEP Students and Use of Accommodations and Adaptations in Large-Scale Assessments

It is crucial to understand current practice in assessing LEP students, both nationally and at the state level. When examining state policies and practices, there are several questions of paramount importance: How many LEP students currently participate in statewide assessment programs, what kinds of special testing conditions or accommodations are provided to enable them to participate, and how are the results for LEP students reported? The following sections present information that examines these questions.

Participation of LEP Students in State Assessment Programs

As discussed in the previous chapter, the State Student Assessment Program provided data on state assessments of students with disabilities and LEP students from their 1994-1995 survey (Bond et al. 1996). According to the survey, 36 states allowed for the exclusion of LEP students. There are many similarities between the findings for students with disabilities

and LEP students. Many states allowed schools to exclude students if the assessment was judged not appropriate for them, for example, if the LEP student does not know enough English to complete the exam successfully. Few states collected data on the numbers of LEP students who were excluded, and few could determine what percentage of the total population of LEP students were excluded from the assessment.

For LEP students, the level of English proficiency and/or the number of years the student has studied English as a second language were the determining factors used for decisions whether the student could participate. As was found for students with disabilities, many states reported that assessment results for LEP students may be eliminated from state, district, and school summaries.

State's Use of Accommodations and Adaptations in the Assessment of LEP Students

When LEP students were included in statewide assessments, the use of accommodations for them varied. Only seven states included LEP students without offering accommodations, and 25 states included them with accommodations. From the SSAP data, the following table (table 3.4) shows the responses of 17 states to the question: "What kinds of testing accommodations do you allow for LEP students?" Compared to the data in table 2.3 on the permissible accommodations for students with disabilities, a much smaller number of accommodations are allowed for LEP students. However, the findings may need to be qualified somewhat because the question asked in the SSAP survey did not focus specifically on LEP students, but was asked more generally about all types of accommodations that were permitted for either IEP or LEP students. Nevertheless, nearly all of the 17 responding states reported allowing the use of separate scheduling and testing settings, small group administrations, and extra time. About half of the responding states allowed audiotaped instructions, multiple/extra testing sessions, simplification of directions, and the use of dictionaries. Only four states reported that they allowed other languages to be used with LEP students, and only three states administered an alternative form of the test. Also note that, similar to the caveat stated for the data in table 2.3, the data in this table are based on the accommodations states permit for LEP students and not on what is actually used.

Table 3.4—Permissible Accommodations for LEP Students

Type of Accommodation	Number of States that Permit Accommodation
Separate Testing Session	17
Flexible Scheduling	15
Small Group Administration	15
Extra Time	14
Simplification of Directions	11
Large Print	10
Other Accommodation	10
Audiotaped Instructions/Questions	9
Multiple/Extra Testing Sessions	9
Use of Dictionaries	9
Braille/Sign Language	8
Word Processor	8
Audiotaped Responses	4
Other Languages	4
Alternative Test	3

SOURCE: The Status of State Student Assessment Programs in the United States (CCSSO/NCREL 1996)

Reporting of Results for LEP Students

Another issue that states are grappling with is what to do with the results for LEP students who participated in an assessment. Some states may report results separately, and other states may not report them at all. Although some states may have scores for LEP student performance, the achievement data for the general population may not include their scores. In this case, the state's policy and use of scores from LEP students becomes very important. For educators and administrators involved in improving the education of LEP students, scores used to monitor the improvement of LEP students and target programs are highly desired, and scores only being used to give a cross-sectional look at student performance are less valuable.

C. Summary of Current and Ongoing Research Activities and Projects

Unfortunately, the knowledge base for the best approaches to use when including and assessing LEP students is not very complete. Although much data are being examined and summarized to answer some of the more vexing questions, and some promising progress is being made in the field, many more questions than answers currently exist on the topic. Many studies are underway to investigate various aspects of accommodating students from special populations in large-scale assessments, and many different organizations are involved in these efforts. NCES is closely following, and in many cases, working with, the

following organizations on projects and activities that focus on the inclusion of LEP students.

NCES Projects and Activities

NCES has initiated and funded an active research program to look into a number of issues associated with the participation of LEP students in large-scale assessments, and NAEP, in particular. Although some of these studies are being done by NAEP contractors, several others are being conducted independently by other researchers and organizations. Current and ongoing research related to the operational aspects of including LEP students in NAEP are discussed in detail in Chapter 5.

Prior to the modifications made to NAEP in 1995, NCES began looking into the issues of including more LEP students in NAEP, and what types of accommodations or adaptations of the assessment were best suited to increase participation. Input was provided by a wide variety of researchers and practitioners knowledgeable of the topic of educating LEP students. The next sections describe these activities.

NCES Conference on Inclusion Guidelines and Accommodations for Limited English Proficient Students in NAEP. NCES convened a conference on assessing LEP students in December 1994, which included a number of experts and researchers who focused on issues of including LEP students in assessments. The primary purposes of the meeting were to inform NCES on possible modifications in NAEP administration procedures to make it more inclusive of LEP students, to develop guidelines for the increased inclusion of LEP students, to examine policies on the reporting of data, to investigate the major technical and implementation issues that might be part of a federal research agenda on inclusion and accommodations, and to plan follow-up research that will help ensure the adoption of appropriate and consistent inclusion and modification strategies (U.S. Department of Education, NCES 1994).

The NCES report *Proceedings from the Conference on Inclusion Guidelines and Accommodations for Limited English Proficient Students in NAEP* (U.S. Department of Education, NCES 1996) summarized the results of this working meeting to provide guidance to staff at NCES on inclusion guidelines and accommodations for limited English proficient students in NAEP. In general, conference participants emphasized the importance of developing a set of guidelines for determining how to include LEP students in NAEP. At the meeting, the group cautioned about assessing LEP students in their native language, since these students may not be literate in their native language. The group also voiced concern that translations should be conducted with great care because of the difficulty in vocabulary that varies from one language to another and because some languages like Spanish have numerous variations in the ways certain words or phrases are stated. Other important issues discussed were:

- The importance of a "standardized" definition of Limited English Proficient.

- The belief that only those LEP students proficient enough in English to participate meaningfully in NAEP should be given the unmodified version of NAEP.

- The importance of developing scoring rubrics and procedures that are appropriate for LEP students, (i.e., that consider their particular linguistic and cultural background).

- The need to develop criteria to determine the best match between the particular characteristics of LEP students and the particular form of assessment.

- The need, in the process of determining which accommodations should be used with LEP students, for the assessment to be aligned with the educational program. For example, LEP students who do not receive any mathematics instruction in Spanish will find it difficult to take a test in Spanish.

- The suggestion that efforts were needed to report outcomes for LEP students by type of accommodation, and that NCES should consider reporting data separately on LEP students who took the standard NAEP assessment with no accommodation, as well as including their results as part of the total U.S. aggregate average.

Although the working meeting raised many issues about how to include LEP students in NAEP and other national assessments, it provided little resolution of the issues. One of the most important issues raised in the meeting was the need for the development of a definition of what constitutes limited English proficiency that could be consistently applied across states and schools. Then, once this definition is available, it could be implemented using appropriate measures to determine if an individual student is LEP and to determine how to assess these students.

The report concluded that inclusion strategies for NAEP will require much more research, development, and field testing before they can be implemented. Conference participants recommended that criteria be established to determine which methods can be field tested now and which require more research and development work. Participants also recommended that an advisory committee be established to provide ongoing advice to NCES on LEP student assessment issues, to review ongoing research, and to make recommendations on research needs. This recommendation led to the establishment of an Education Department working group on students with disabilities and LEP student issues prior to the 1995 NAEP field test.

Working Paper on Assessing Students with Disabilities and Limited English Proficiency. As discussed in Chapter 2, the NCES working paper *Assessing Students with Disabilities and Limited English Proficiency* (Houser 1995) was disseminated in 1995 to describe the current state of affairs as it related to policies in place at NCES on the exclusion of some students from NCES assessments and surveys. The following concerns were expressed as

they related to issues of data validity, assessment modifications, and the inclusion of LEP students.

- If knowledge about a subject area were the primary policy interest of the study, LEP students could be assessed in their native language. However, since the primary policy interest behind many education assessments is the ability of U.S. students to compete economically, and because there are economic advantages to individuals in the U.S. that result from fluency in English, assessing LEP students in their native language is generally considered inappropriate.

- Because Spanish is by far the most common language spoken by LEP students, NCES would be unlikely to obtain sufficient sample sizes within a randomly drawn sample in order to report separate test scores for languages other than Spanish.

- Since Spanish is spoken in many different dialects, NCES might not be able to administer just one version of a Spanish test.

- Students may not always be literate in the language spoken at home and some languages are only spoken and not written. As a result, some tests may be difficult to administer to LEP students even if the tests are translated into a student's native language.

- Test scores may need to be reported separately for each language, because some psychometricians believe that assessments conducted in different languages are not psychometrically equivalent. This is done for the SAT.

The Office of Bilingual Education and Minority Languages Affairs

NCES has received valuable guidance from the Office of Bilingual Education and Minority Languages Affairs on the conduct of assessment activities related to language minority students. OBEMLA funds a variety of studies and activities related to the education of LEP students. These activities include grants to State Education Agencies, the Academic Excellence initiative, support services, bilingual fellowships, benchmark studies on bilingual education and school reform, and the National Clearinghouse for Bilingual Education (NCBE).

The NCBE provides a valuable resource to OBEMLA. Its charge is to collect, analyze, synthesize, and disseminate information related to the education of linguistically and culturally diverse learners in the United States. It is operated by the George Washington University Center for the Study of Language Education. The NCBE serves as a clearinghouse for information, including much information on studies of the assessment of language minority students and on LEP student performance. It provides information services through a variety of mechanisms (such as online services, telephone reference and referral service). As part of the U.S. Department of Education's technical assistance and

information network, NCBE also works with other service providers to provide access to high quality information to help states and local school districts develop programs and implement strategies for helping all students work toward high academic standards. More details on the services provided by the NCBE are provided in Appendix C of this report.

The National Academy of Education's Studies of Assessing LEP Students

As was described in the previous chapter, several studies have been carried out as part of the National Academy of Education's evaluation of NAEP. In addition to focusing on issues related to the inclusion and accommodation of students with disabilities in the assessment, the NAE has also looked into issues that deal with LEP students.

- The NAE has produced several papers commissioned by NCES on issues associated with the assessment of LEP students. The first, "Issues in the Development of Spanish-Language Versions of the National Assessment of Educational Progress" (Secada 1994), addressed the feasibility of conducting NAEP in Spanish. The paper concluded that NCES should conduct a pilot study to examine the technical implications of a Spanish-language assessment and NCES should try to design NAEP so that it can capture a broader range of students than it currently is doing.

- Another paper, "A Study of Eligibility Exclusions and Sampling: 1992 Trial State Assessment" (Spencer 1994), examined the exclusion of IEP and LEP students from the assessment. The author concluded that, for LEP students, the cost-benefit analysis implicit in the exclusion decision needed to be examined since the decision to exclude some students may be based on the difficulty in assessing them.

- In the Third Report to Congress on the evaluation of the NAEP Trial State Assessment (NAE 1993), the NAE panel recommended that serious consideration should be given to the development of non-English versions of NAEP. They pointed out that the exclusion rates for the 1992 Trial State Assessment varied significantly state by state, and that these differences should not be surprising given the ethnic compositions of the states. This was of concern since these state-by-state variations could impact the ability to make valid cross-state comparisons using NAEP data. The panel also recommended that NAEP conduct pilot studies in locales where significant proportions of Spanish-speaking students lived, and that these studies carefully examine whether NAEP results vary as a function of administering a Spanish-language version of NAEP to Spanish-speaking students.

- In a follow-up study for the NAE evaluation, researchers at AIR (Stancavage et al. 1996) focused on fourth-grade students excluded from the 1994 Trial State Assessment in reading to determine the assessibility of the excluded students and the types of adaptations that would be needed to include them in future assessments. Research was also done to obtain more detail about the implementation of the exclusion decision process in order to improve the directions provided to local school sites. The study included an analysis of questionnaire data on all LEP

students, whether or not they were excluded from the assessment, as well as teacher interviews. The results showed that the percentage of time per week spent in a special language program, such as bilingual, and the exclusion of these students from state, district, or other grade-level standardized tests were positively associated with exclusion. LEP students who spent more time in a bilingual program or who were excluded from other tests were more likely to be excluded from NAEP than students who were not. Students who spent more years in a special language program and whose teachers gave higher estimates of their functional grade level for writing were less likely to be excluded.

- In the Fourth Report to Congress by the NAE panel (NAE 1996), findings and recommendations from the second study of the assessability and exclusions among LEP students were provided. The panel found that a high proportion of LEP students would have been assessable on the current NAEP instrument. They suggested the possibility that "achievement for some LEP students may be underestimated by NAEP because of linguistic demands of the assessment that are not essential to the competencies the assessment is attempting to measure." The panel also found that teachers were quite liberal in recommending the use of accommodations, adaptations, or exclusions for their students. They noted that, if teachers' recommendations had been followed, about half as many (30 percent instead of 63 percent) of the LEP students in their study would have been assessed under standard conditions in the 1994 TSA. Lastly, the NAE evaluation reported "perhaps the most disturbing finding from the LEP study was the significant fraction of these LEP students who were excluded from the assessment despite the fact that more than three-quarters had four or more years in English-speaking schools" (NAE 1996, p.69). The panel recommended that NCES continue its efforts to identify appropriate adaptations or accommodations and permit the inclusion of larger proportions of LEP students in the assessments.

The Prospects Study

The Prospects Study was mandated by Congress to evaluate the Chapter 1 (now Title I) program. An OBEMLA monetary supplement expanded the study by funding a sub-study to examine LEP students in the Title I program. Title I, the largest single federal elementary education program, provides funds to school districts to provide supplemental instruction to low-achieving students in low-income schools. Instructional areas supported by Title I include reading, mathematics, language arts, English as a second language (ESL), and bilingual education. Congress made it clear in the 1994 reauthorization of Title I that LEP students are eligible for Title I services. Although the Prospects Study does not provide detailed data on the nature of bilingual education or English as a second language instruction received by these students, this Congressionally mandated study includes nationally representative data on students who were in the first, third, or seventh grade at the beginning of this five-year project period. It includes a supplemental sample of first and third graders in schools with high concentrations of LEP students. Information is collected annually from students, parents, teachers, and principals.

In the reports on *PROSPECTS: The Congressionally Mandated Study of Educational Growth and Opportunity* (U.S. Department of Education 1993, 1996), a number of findings were presented. The interim report was based on data for first and third graders in school year 1991–92. Subsequent reports on the Prospects Study will examine student outcomes over time. The overall results found that the academic performance of LEP students generally lagged behind other elementary school students, as measured by grades, retention in grade, teacher judgments of student ability, and standardized tests. Eighty to 90 percent of LEP students receive some form of supplementary education from federal, state, or local sources. However, 30 percent do not receive ESL or bilingual education from any source. In schools receiving Title I funding, about one out of five LEP first and third graders with low performance in mathematics or reading receive no supplemental instruction in those areas.

Nationally, children reared in a setting in which a language other than English is commonly used comprise about 16 percent of all students in the first and third grades. Among these language minority students, about 40 percent are classified as LEP. While these students vary widely, many are particularly disadvantaged—they come from very poor families and typically live in communities (mostly urban) with high concentrations of poverty. Often their parents have only limited proficiency in English and may rarely speak English at home, are more likely to be employed in low-income jobs, and have low levels of formal education.

Research of the Stanford Working Group

In 1994, a diverse group of experts on the education of LEP students developed a consensus document with recommendations regarding state-level assessments and Goals 2000. In the document *For All Students: Limited English Proficient Students and Goals 2000* (August et al. 1994), the following recommendations were presented:

- If LEP students are not assessed then no one can really be held accountable for what these students know and can do in important content areas. Therefore, states need to develop performance assessments that are appropriate for LEP students.

- LEP students who are instructed in their native language should be assessed in that language. The native language assessments should parallel content assessments and performance standards in English.

- Modifications in assessments and assessment procedures should be encouraged to enable LEP students to take content assessments in English. These modifications might entail: altering the procedures used to administer the assessments, modifying the assessment itself so it is more comprehensible to LEP students, using alternative assessments, and employing computer-assisted assessments that are tailored to the language needs and content knowledge of LEP students.

- Until the psychometric issues underlying these new assessments have been addressed, and until mechanisms to ensure opportunities to learn have been fully implemented, these assessments should not be used in high stakes situations.

Following meetings with NCES, a working paper was prepared to help develop strategies for improved LEP student inclusion (Hakuta and Valdes 1994). The primary goal was to make appropriate and consistent decisions about assessing LEP students, and a secondary goal was to maximize inclusion. The paper focused on NAEP, particularly the state component, and identified several options that could help to develop ways to include LEP students in the assessment. In brief, the researchers suggested two general approaches for studying inclusion: (1) empirically test various options to maximize the number of students who are offered appropriate and valid accommodations and conduct more research to increase the validity of options that may not be as appropriate or valid, and (2) consider only options that are realistic in the context of policy and NAEP. Their proposal recommended a study that would randomly assign test conditions to students in order to examine a variety of options. These included native language assessments in Spanish, assessments in English, bilingual side-by-side assessments, modified administration procedures using unmodified items, and modified administration procedures using linguistically modified items.

Activities and Plans of the SCASS Assessing LEP Students Project

The purpose of this SCASS group, a developmental consortium of states (see Chapter 2 for general information on CCSSO's SCASS consortia) is to encourage the appropriate assessment of students learning English in large-scale and other assessments conducted at the state and local levels. The group has developed plans for several types of activities:

- Improved assessments of language proficiency for use in the selection and placement of students in language classes.

- The development of content assessments, such as science, that are appropriate for students learning English.

- Conduct basic research on effective English language learning to use as a basis for the development of appropriate assessments at the classroom level and for use in large-scale assessment contexts.

- Work on the development of accommodations for LEP students.

The SCASS group has been meeting with national experts concerned with issues in the instruction and assessment of LEP students, and has been discussing ways of determining exemplary classroom instruction related to the use of English. They are working with experts and teachers in the field to articulate the benchmark behaviors of English proficiency within content-related activities. These activities will relate to the content standards and the types of constructive tasks required of students on national and state performance assessments.

During 1996, the SCASS LEP project developed a set of guidelines for scoring the responses of LEP students to constructed-response items. This was part of a larger effort to develop guidelines for effective language learning and the assessment of LEP students on statewide content standards for language and content areas such as mathematics, science, and social studies. Plans are to continue work on the development of measures of English proficiency for LEP students. The scoring guidelines were developed in the area of mathematics and are used to sensitize scorers of constructed-response items to the ways in which LEP students may experience challenges in completing these types of performance assessments. Plans are to conduct a pilot study of the guidelines during school year 1996–97 to determine whether more accurate scoring for content knowledge results is possible.

CRESST Studies of Language Issues

The National Center for Research on Evaluation, Standards, and Student Testing (CRESST) is a partnership of UCLA, the University of Colorado at Boulder, Stanford University, the RAND Corporation, the University of Pittsburgh, and the University of California, Santa Barbara. CRESST studies focus on the assessment of education quality and their studies have encompassed a wide variety of topics in the area. Of particular interest for this report, researchers at CRESST have conducted, or are in the process of conducting, several studies that examine the linguistic features of NAEP test items. The goal of these studies is to identify linguistic features in mathematics items that may affect the performance of students with language backgrounds other than English.

Language Background and Performance on NAEP. Results from CRESST's first study on language background and linguistic features (Abedi et al. 1995) found lower mathematics proficiency scores for students who predominantly spoke a language other than English at home. These students also performed less well on more lengthy types of test items, as well as on linguistically complex items, which were defined by non-familiarity or frequency of non-math vocabulary, voice of the verb phrase, length of nominals, conditional clauses, relative clauses, question phrases, and abstract or impersonal presentations. These students also had much higher percentages of items omitted and not reached than did students who only spoke English at home.

In investigating the role of linguistic complexity in students' mathematics performance on both the original NAEP items and those revised to be linguistically simpler, the study found that students with a wide range of different mathematics skill levels reported a strong preference for the revised and simplified items. This was true for both native and non-native speakers of English.

In summary, the study found that language minority students exhibited much lower mathematics performance than other students. This difference in performance could not be explained by differences in socioeconomic status or other background variables. Therefore, the authors concluded that language must be a very important component, and that ability to use English and performance on the assessment are confounded. The authors also found that

revising mathematics items to make them less linguistically complex helped to improve some students' performance, particularly those in low- and average-level mathematics classes. Thus, they also concluded that the language of mathematics items may disproportionately impact the scores of less language-proficient students, whether they are native or non-native speakers.

Research on Language Background and the Validity of Accommodations and Modifications. Currently, CRESST is involved in additional studies to examine further linguistic issues and the performance of LEP students. Two new research efforts proposed by CRESST researchers are underway. They focus on the validity of accommodations and modifications in assessments. These studies replicate certain aspects of the original Abedi et al. (1995) study described above, but with two major changes—a specific focus on LEP students and an attempt to obtain a measure of English proficiency that will help relate test performance to language in a more precise manner. As proposed, the primary goal of this research is to produce a continuum of accommodations and modifications that may be appropriate and feasible for use in NAEP. A second goal is to improve the procedures for matching students to assessment modifications for students whose primary language is Spanish. These studies are using an experimental research design to compare accommodated or modified versions of the assessment items with appropriate comparison groups. The first study provides data on language background, English language proficiency, and subject area (mathematics) performance. The second study investigates the role of native language proficiency and performance on a wide variety of assessment accommodations, including linguistic simplifications of the test items.

Research Framework for Investigating Accommodations. Lorrie Shepard, of the University of Colorado at Boulder, is involved in another study that is being conducted by CRESST. She has developed a research framework for investigating accommodations for language minority students. As discussed at the 1996 CRESST Assessment Conference at UCLA (Shepard 1996), a preliminary framework is proposed to guide a research agenda for the assessment of language minority students. The framework has two dimensions—the domain of proficiency to be assessed (knowledge of subject matter, native language and literacy, or English language and literacy) and assessment purpose (instruction or accountability). Performance, as measured by an assessment, is viewed as a continuum related to language acquisition, with an emphasis on issues related to logistics and adaptations, validity and fairness, and consequences of the assessment (Shepard 1993).

Because the resolution of both practical and theoretical problems differs depending on what is being measured and how the results will be used, Shepard identified four research questions:

1. How should a performance continuum be conceptualized and with what benchmarks of increasing proficiency?

2. What logistical problems or adaptation needs must be addressed to make it feasible to assess language minority students?

3. Do assessments with adaptations ensure valid and fair demonstrations of learning for language minority students?

4. What are the consequences of assessment use?

Plans for this study include the gathering of detailed information at a small scale level, via interviews and cognitive workshops, and a focus on data that closely reflect students' thinking processes.

Research Activities of The George Washington University Center for Equity and Excellence in Education

In 1995, The George Washington University Center for Equity and Excellence (GWU/CEEE) in Education initiated a national project to develop a set of tools for policymakers, educators, and community members seeking sound research-based models and strategies to design or evaluate services for limited English proficient students. The *Promoting Excellence Project* identified six principles that represent "best practice" in an optimal learning environment for limited English students. The principles describe ideal conditions in which LEP students succeed academically. The principles are a tool that educators can use to facilitate the inclusion of LEP students in local reform initiatives supported by *Goals 2000* and the *Improving America's Schools Act.*

The six research-based principles found in *Promoting Excellence: Ensuring Academic Success for LEP Students* (GWU/CEEE 1996) address a range of issues, including equity, teaching and learning, and appropriate assessments. They challenge educators:

- to hold LEP students to high expectations (Principle 1),

- to teach LEP students English language skills comparable to those of native English speakers (Principle 2),

- to provide opportunities for LEP students to study challenging content and to reach the same performance standards as monolingual English students (Principle 3),

- to utilize instruction that builds on prior educational background and that is sensitive to the language acquisition stage of students (Principle 4),

- to evaluate LEP students with valid and reliable assessments aligned with state and local standards (Principle 5), and

- to recognize that the academic success of LEP students is a responsibility shared by all educators, the family and the community (Principle 6).

Principle 5 supports the use of valid and reliable assessments aligned with state and local standards to evaluate LEP students on the same content and performance standards as monolingual English students. Among the many recommendations made for meeting this principle, the indicators support: (1) the use of test accommodations, (2) the use of flexible assessment systems that incorporate multiple assessment strategies and that do not solely rely on English language norm-referenced tests, (3) consideration of the whole profile of a student, including language/cultural background and native language literacy, when monitoring an individual student's academic progress and when evaluating an instructional program, and (4) regularly reviewing whether the assessments are being used appropriately (e.g., for purposes for which they have been validated) with LEP students.

Study of Policies and Practices for the Inclusion of LEP Students. A recent study from the GWU/CEEE summarized findings from a national survey of state assessment directors (Rivera, et. al. 1997). The report, *Statewide Assessment Programs: Policies and Practices for the Inclusion of LEP Students*, presents findings on state assessment policies and practices, identifying key issues affecting the measurement of LEP students' academic progress. Relevant data from CCSSO/NCREL's survey were analyzed with data from the GWU/CEEE survey of state assessment directors. The GWU/CEEE survey, *Policy and Practices 1993-1994 Statewide Assessments: Participation of LEP Students*, provided data in five key areas: (1) the prevalence of statewide assessment programs, (2) states' propensity to exempt LEP students from statewide assessment programs, (3) the use of test modifications in statewide assessment instruments, (4) the use of assessments in languages other than English, and (5) the use of statewide assessments for student and school accountability. Among the report's most salient findings were:

- Nearly every state (48) had an assessment program underway, and most (44) allowed exemptions for LEP students, with the most common criteria for exemptions based on English language proficiency level (61%) and time in the United States or school district (45%).

- Few states (33%) reported the actual number of LEP students assessed in their state.

- About one-half (52%) of the states allowed one or more test modifications for LEP students.

- The most frequently allowed modifications were extra time (81%), small group administration (74%), flexible scheduling (63%), simplified or clarified directions in English (56%), and use of dictionaries or word lists (52%).

- Only four states translated tests or developed tests in languages other than English.

Policies and Practices for the Use of High School Exit Exams with LEP Students. Data from the GWU/CEEE survey of state assessment directors, together with those from the SSAP CCSSO/NCREL survey, were used to document assessment policies for LEP students in states that require a test for high school graduation. Findings from the report *High School Graduation Testing: Policies and Practices in the Assessment of Limited English Proficient Students* (Rivera and Vincent, in press) indicated that in school year 1993-1994, 17 states required that students pass one or more content area tests to receive a standard high school diploma; state policies were limited in their consideration of LEP students; and states do not generally collect data on the number of LEP students who take exit exams. Deferring LEP students from the first administration of the test and permitting the use of test modifications (e.g., allowing extra time), were two common practices used by states in addressing the assessment of LEP students.

The authors suggest that, while the use of test modifications offered some benefits in making high school graduation tests more accessible to students, test modifications were limited in their scope and frequency of use and did not provide students with appropriate access to the test, even for the limited number of students who were provided modified tests. Rivera and Vincent recommend that states: (1) develop better policies for dealing with LEP students in their testing programs, (2) improve their data collection systems to track LEP students, (3) find alternatives to deferring LEP students from taking the first administration of the required high school graduation test, (4) use test modifications judiciously, (5) administer native language tests when appropriate, and (6) rely on alternative assessments that lessen the language load. The authors conclude that, although the survey data captured policies and practices at a specific point in time, the findings represent trends in the testing of LEP students. Plans are to conduct a follow-up survey of these testing practices to collect more current data and to get another look at recent trends.

Report from the National Research Council

The NRC recently completed an extensive study of LEP students, which is summarized in a report, titled *Improving Schooling for Language Minority Children: A Research Agenda,* (August and Hakuta 1997). In 1994, a Committee on Developing a Research Agenda on the Education of Limited English Proficient and Bilingual Students was formed by the NRC to study the issues involved in the education of these students, review the methodologies used in this area, and make recommendations regarding promising research activities in order to improve policy and practice in the field. As stated, the purpose of the report is to contribute to the development of a knowledge base in the education of students who are not fully proficient in English. To do this, the report provides a review of the state of knowledge and identify a research agenda that will address key gaps in present knowledge on the topic. Among the topics covered in the report are student assessment, program evaluation, and estimating population parameters to obtain education statistics from surveys such as NAEP and other large-scale surveys conducted by the U.S. Department of Education.

D. Summary

Most large-scale assessments are meant to assess the knowledge and skills of all students. Students whose native language is not English, however, are often excluded from participation in these assessments. Increasing the participation of these students entails a number of methodological and resource issues. A number of these were discussed in this chapter.

According to OBEMLA, in 1995 there were approximately 3.2 million limited English proficient students in the United States. This was an increase of 4.8 percent from the previous year. Other data sources found about two-thirds of LEP students were concentrated in five states: California, Texas, New York, Florida, and Illinois; about three-fourths of them are Spanish-speaking; and LEP students tended to be more concentrated in the lower grades.

Limited English proficient students may be assessed by states and school districts under a variety of practices. Guidelines used to decide whether to include LEP students in assessments vary widely among states. Some guidelines are based on the number of years students have been exposed to instruction in English, while others are based on the judgments of school staff.

Policies and practices for offering testing accommodations to remove barriers to participation for LEP students also differ. The most frequent accommodations offered include separate or small-group testing sessions for LEP students, flexible scheduling arrangements, allowing extra time to complete the assessment, simplification of directions, and provision of word lists or glossaries in the student's native language. Translations of assessments into students' native languages is not done as often, usually because translations can be expensive and languages such as Spanish have a variety of dialects with which a particular student may not be familiar. Some students who are not proficient in their native language, as well as students who are not instructed in their native language, probably would not benefit from a translated version. In addition, for some academic subjects, particularly reading and writing, students are assessed for their ability to perform in the English language. Many educators believe, further, that a primary goal should be to teach English-language skills to limited English proficient students, so that they can function in a predominantly English-language society.

The assessment of LEP students has been the subject of a wide range of research in recent years. A key issue is how to assess a student's content knowledge when it may not be readily expressed in English. Another important issue, recommended by several different researchers, is the need to include LEP students in the design and piloting of assessments. Aside from the translation issues discussed above, research is being conducted into: the effects of native language literacy on performance in English-language assessments; the use of simplified versions of tests in English; development of scoring rubrics and administration procedures sensitive to the linguistic patterns and cultural orientations of LEP students; whether more LEP students could participate in assessments (and how to accomplish that

objective); which testing accommodations would be most appropriate; and the validity of assessment results for students assessed under accommodated conditions, as compared to those assessed under standard conditions.

Policy makers are in need of more accurate data about limited English proficient students. The lack of a standard definition of LEP, common to all states, is also an issue. Many states do not provide data on the number of LEP students in their schools. Many also do not provide information on the number of such students excluded from assessments. Better data would allow for a more accurate understanding of the extent and nature of the situation, as well as better estimation of the resources needed to include these students in large-scale assessments.

In the next chapter, the steps taken to increase the participation of both students with disabilities and LEP students in NAEP are discussed.

Chapter

4

Approaches and Procedures Implemented by NAEP in Recent Years to Increase the Participation of Students with Disabilities and Limited English Proficient Students

A. 1995 NAEP Field Test
   - NAEP's approach prior to 1995
   - Degree of exclusion of students identified as IEP or LEP
   - Exclusion/inclusion criteria used in NAEP
   - Field test of revised inclusion criteria and new accommodations procedures
   - Findings from the NAEP field test

B. The Puerto Rico Special Assessment
   - Factors leading to a special NAEP-like assessment in Puerto Rico
   - Design and operational issues for the Puerto Rico Assessment of Educational Progress
   - Findings from the Puerto Rico Assessment

C. 1996 NAEP assessment
   - Special sample design and implementation of new inclusion criteria and accommodations
   - Reporting 1996 results

D. Summary

As discussed in Chapter 1, NCES is currently in the process of examining and implementing new procedures for NAEP to maximize the representativeness of students included in the assessment. In recent years, NCES has been involved in a number of efforts to develop more inclusive assessments and to enhance overall participation by including more students with disabilities and LEP students. These efforts have included making adaptations of NAEP, such as the development of Spanish-language assessments, and providing various types of accommodations. Additional work has focused on reviewing and revising the criteria used to exclude or include students in NAEP.

In this chapter, the focus is on summarizing the approaches and procedures implemented by NAEP in recent years to increase the participation of students with disabilities and LEP students. An overview of activities recently conducted or currently underway will be presented, with inclusion-related activities from three key assessments highlighted. First, the plans leading to modifications of procedures in the 1995 NAEP field test toward the goal of increasing the participation of students in the national assessment, as well as the findings from the field test, are summarized. Second, a summary of the special NAEP-like assessment project conducted in Puerto Rico is discussed. Finally, the implementation of new procedures operational in the 1996 NAEP and the special design used to study the effects of new inclusion criteria and providing accommodations are discussed.

A. 1995 NAEP Field Test

In 1994, NCES began implementing plans to increase the participation of both students with disabilities and LEP students in NAEP. Along with the data that were made available from the Puerto Rico Assessment of Educational Progress experience in 1994 (discussed in the next section), NCES also received input from a variety of organizations and researchers on how to make NAEP more inclusive, with the goal of implementing new procedures for the 1996 national assessment. To achieve this goal, plans were developed to try out new procedures in the 1995 field test, including revised inclusion criteria and a variety of accommodations.

NAEP's Approach Prior to 1995

In the past, students with disabilities and limited English proficient students were often excluded from NAEP for several reasons—the Individualized Education Program specified the student should not be assessed, no test accommodations or adaptations were available that met the specific needs or requirements of the IEP required by law for students with certain disabilities, state and local policies had been designed or implemented to identify students with disabilities and LEP students and exclude them from testing based on certain criteria, or school staff may have believed these students were unable to participate meaningfully. In order to standardize NAEP procedures, NAEP had previously developed policy guidelines for including students with disabilities or LEP students, but some were excluded, particularly those with profound disabilities or students who would have required accommodations to the testing procedures.

Degree of Exclusion of Students Identified as IEP or LEP

In previous years, about half or more of the students identified as IEP or LEP were excluded from the NAEP assessments (see table 4.1).[7] The overall percentages of students excluded, because in the judgment of school staff they could not participate meaningfully, have been fairly steady over time, with about five percent of all students excluded due to IEPs and one

[7] In this section, the term IEP is being used (instead of students with disabilities) because this terminology was in effect at the time and was used by NAEP to report the results referenced here. In earlier assessments, the IEP classification encompassed students with disabilities.

percent excluded due to LEP. While the percentage of students excluded has remained small in relation to the total population, these students make up a relatively large portion of the special education and LEP populations. Decisions to exclude students from the assessment were made at the local level based on specific criteria used with NAEP (discussed further in the next section). Exclusion was further impacted by NAEP's previous policy of not offering accommodations and adaptations, such as extended testing time, Braille versions, or assessment instruments in other languages.

Table 4.1—NAEP Inclusion Rates for IEP and LEP Students, by Grade: 1992–1994

	Percentage of Students Identified IEP	Percentage of Students with IEPs Assessed	Percentage of Students Identified LEP	Percentage of LEP Students Assessed
Grade 4				
1992	9	33	4	25
1994	12	50	6	50
Grade 8				
1992	9	44	3	33
1994	13	38	4	50
Grade 12				
1992	5	20	2	50
1994	11	36	3	66

As discussed in the previous chapters of this report, evaluations of NAEP conducted by the National Academy of Education (NAE 1993, 1996) found that many of the students with disabilities and LEP students who had been excluded from NAEP were, in fact, capable of participating in the assessment. This was particularly true if certain types of adaptations and accommodations could have been offered for the assessment. The NAE also recommended that NAEP develop better criteria to promote inclusion, rather than exclusion, of students. During this same time period, other groups also were involved in examining the issues related to including more students in assessments. These groups provided additional valuable input to NAEP on ways to improve the inclusion and accommodations process (NCEO SR17 1994; August and McArthur 1996). In response to these concerns, NAEP explored ways to increase the inclusion rates even further. To meet these objectives, plans were made to try out new approaches toward inclusiveness in preparation for the 1996 assessment.

Exclusion/Inclusion Criteria used in NAEP

The exclusion or inclusion criteria used in NAEP have changed somewhat over time. Prior to 1995, the procedures used by NAEP to determine who can participate in the assessment were based on criteria for excluding students. Beginning with the 1995 NAEP field test, the

criteria were revised with the intention of making appropriate and consistent decisions about the inclusion of students with disabilities and LEP students. Other changes were made to make a better link between subject areas of instruction and the assessment and in response to concerns about the language used in the old criteria. The old and new criteria are presented in table 4.2.

Table 4.2—Old and New Exclusion or Inclusion Rules for NAEP

	Old (1990–1994)	New (1995–1996)
	Exclude If:	**Include If:**
Students with Disabilities	Student is mainstreamed less than 50 percent of the time in academic subjects and is judged incapable of participating meaningfully in the assessment, OR IEP team or equivalent group determine that the student is incapable of participating meaningfully in the assessment.	Student has an IEP, unless the IEP team or equivalent group determine that the student cannot participate, or if the student's cognitive functioning is so severely impaired that he or she cannot participate, even with accommodations.
Students with Limited English Proficiency	Student is native speaker of a language other than English, AND Enrolled in an English-speaking school (not including bilingual education program) for less than two years, AND Judged to be incapable of taking part in the assessment.	Student has received academic instruction primarily in English for at least three years, OR Student has received academic instruction in English for less than three years, if school staff determine that the student is capable of participating in the assessment in English, OR Student, whose native language is Spanish, has received academic instruction in English for less than three years, if school staff determine that the student is capable of participating in the assessment in Spanish (if available).

In all years, schools were instructed that students with disabilities and LEP students should be assessed if, in the judgment of school staff, they were capable of taking the assessment, and that, in cases of doubt, the school should err on the side of inclusion.

Field Test of Revised Inclusion Criteria and New Accommodations Procedures

For the 1995 field test, NAEP procedures used in previous assessments were revised with the intent of meeting two goals—including more students in the assessment and making inclusion policies more consistent. Changes were made to the criteria for including students in the assessment and accommodations were offered. Two special studies were undertaken in the 1995 field test to study how more students with disabilities and LEP students might

participate in the assessment, and how inclusion policies might be made more consistent from one area of the country to another. The LEP special study explored the impact of changing the criteria for inclusion to be more consistent with current state practices, and whether administering the assessment in Spanish would be both feasible and valid. The special study of students with disabilities also focused on the impact of changed criteria, and on the feasibility and validity of providing accommodations such as large print or Braille booklets where needed.

In preparation for the 1996 assessment, NAEP field tested in 1995 the new inclusion criteria for participation and the use of various accommodations and adaptations for the mathematics assessment. The field test also included a tryout of new science items, but no accommodations or adaptations were tested for this subject. For mathematics, accommodations were made available for students with disabilities if they were part of the student's normal testing procedure, as specified in the student's IEP.

Students with disabilities were provided a variety of accommodations in administration procedures, if these accommodations were part of the student's normal testing procedure. These included:

- Extra testing time
- Multiple sessions
- Individual or small-group administrations
- Allowing a facilitator to read directions, items, and/or interpret diagrams or graphs
- Allowing students to give answers orally, use sign language, or point to the response option
- Allowing students to give answers using a special mechanical apparatus, (e.g., a tape recorder, Braille typewriter, computer, template, special writing tool, etc.)

In addition, the assessment adaptations for students with disabilities included:

- Provision of large-print booklets and large-face calculators
- Provision of Braille booklets and talking calculators

LEP students were provided many of the same accommodations in administration procedures as listed above for students with disabilities (e.g., extra testing time, modifications in the administration of sessions, facilitation in the reading of directions). Accommodation and adaptation strategies for LEP students (provided for mathematics only) also included the availability of:

- Spanish-English bilingual assessment booklets, with items in different languages presented on facing pages

- Spanish-only assessment booklets

The special study of students with disabilities was conducted at grades four and eight. One mathematics test booklet was adapted in each grade. At grade four, a large-print version was created and at grade eight a large-print and a Braille version were offered. Students using these booklets, and students with disabilities generally, were also provided with accommodations in administration procedures listed above if they were part of the student's normal classroom testing procedure. Most students using these accommodations did so in a one-on-one setting with their normal facilitators. These assessments were untimed, but students were asked to move on to the next section if they took more than three times the usual time limit.

The LEP Special Study was conducted to determine whether a bilingual or Spanish-language version of mathematics could be administered, and whether it would be possible to scale data from these booklets to fit the NAEP scale. This study was also conducted at grades four and eight. Conducted by an administrator fluent in Spanish and English, instructions and background questions typically read aloud in English were read aloud in Spanish. In bilingual sessions, students were told to answer questions in the language they preferred, and were given additional time. In the Spanish-only sessions, students were given the same time limits used in the regular NAEP assessment in English.

Findings from the NAEP Field Test

In the report, *NAEP Inclusion Criteria and Testing Accommodations: Findings from the NAEP 1995 Field Test in Mathematics* (Anderson, Jenkins, and Miller 1996), the procedures used in the studies about the inclusion of students with disabilities and LEP students are described in detail, and their results presented. The design for these studies included augmenting the original sample of students with disabilities for the field test with students from schools for the visually impaired to help obtain a sufficient sample size for statistically reliable results. The sample of LEP students also was augmented to include additional numbers of LEP Spanish-speaking students to aid in the analysis of the results. Each booklet was administered to at least 750 students. However, the authors cautioned that, when interpreting the results from the field test, it was important to note that the field test samples were smaller and less representative than the samples in the full-scale NAEP assessments. In addition, the field test sample was not a random sample of the student population, and thus, the results from the field test were not representative of all U.S. students.[8]

[8] It is also important to note that field test samples are not as carefully controlled for school non-response as are operational (national) samples. Also, the mix of subjects assessed in any given year may affect inclusion rates. Some subjects may be more amenable to providing accommodations than others (for example, the NAEP assessment in reading cannot be translated because it is a test of knowledge of reading in English).

Study of LEP students. As part of the analysis and review for the LEP student study, the Spanish and English versions of the items were examined to identify those with large differences in performance between the groups of students taking the items. The magnitude of possible translation differences were determined by comparing the empirical item response plots for the Spanish versions of the items with the theoretical response functions for the English versions. Several examples of items with problematic translations, confusing language, and other instances that presented an increased potential for misunderstanding by the LEP students taking the Spanish versions were identified.

Response patterns for students taking the bilingual booklets were also examined. Although students were instructed to answer either the English or Spanish version, but not both, some students did not follow these directions. In general, most students followed directions, with a vast majority of them (80–90 percent) answering only the Spanish versions of the items, a minority answering only in English versions (6–16 percent), and very few students answering some Spanish and some English versions (1–9 percent).

Additional analyses of the field test investigated the comparability of the LEP items to the English-only items. This was examined by reviewing two types of data—classical item statistics (i.e., average item score and item-total correlations) and Item Response Theory (IRT) functions comparing the Spanish and English versions. Results from the item analysis indicated that the translated versions of some items may not be parallel in measurement properties to the English versions. This conclusion was based on the large percentages of Spanish items that were found to have poor item statistics and were dissimilar to the statistics for the English versions. In addition, researchers found that LEP students responded to about two-thirds of the items in the Spanish blocks in different ways, based on the item-level statistics. Results from the IRT analysis were less clear because of the small sample size of the data.

Study of students with disabilities. For the study of students with disabilities, analyses similar to those conducted for the LEP items were done. Comparisons of item statistics were made between the samples of students with disabilities and students without disabilities. For each of the grades, it was found that the assessment was harder and less discriminating for the subsample of students with disabilities, with most of the items having lower percent correct statistics, and a majority of the items having smaller item-total correlations. In addition, higher rates of omission were observed for students with disabilities.

The findings based on analyses of descriptive item statistics showed that, in general, students with disabilities responded to many of the items differently than did students in the full sample. The differences in item-total correlations and omit rates indicated that many of the items may be measuring differently for students with disabilities than for the entire group. Again, analyses of IRT data were limited in providing clear information regarding the psychometric issues concerning the scalability and comparability of the data, probably due to the small size of the sample of students with disabilities who participated in the field test.

Based on the overall findings from the studies of LEP students and students with disabilities that were conducted as part of the 1995 field test, the authors summarized:

- Accommodations strategies and procedures like those implemented in the 1995 field test would allow for the inclusion of more students in the national assessment.

- Use of accommodations may have an effect on trend measurement, which will be difficult to evaluate. The authors suggested that a bridge or linking study, which would compare a sample of students using previous practices to a sample using revised practices, be conducted to help determine the effect of the new inclusion practices on NAEP.

- The effect of new inclusion criteria is not likely to be as pronounced as that of offering accommodations.

- Decisions as to how best to use the results of students tested with accommodations will need to be addressed.

In many ways, the findings from the special studies of the 1995 field test appeared to be encouraging. Plans were made for the operational use of accommodations strategies in the 1996 assessment. Plans also were developed to evaluate the effects of these changes on the ability to measure trends, as well as to examine the issues related to the reporting of results. Details of these studies are discussed later in this chapter.

Overall, two issues were the primary focus of the investigation of using new procedures and revised inclusion criteria in the field test. One was the impact on student participation and the other was the impact on student achievement. As discussed, the findings showed that the new procedures could be implemented successfully in the 1996 national assessment. Some students with disabilities and LEP students who would not have participated under previous assessment conditions were able to participate in the field test. It is possible, however, that some students who may not have needed accommodations were assessed with accommodations. This raises important implications for validity and trend measurement. The preliminary analysis of the test items and student performance for both the students with disabilities and LEP samples of students included and assessed under nonstandard conditions (i.e., with accommodations or adaptations) in the field test did not clearly indicate if the results would be comparable to those from other students. Since the lack of clear findings may have been due to the smaller sample sizes used in the field test and the data were not generalizable to those from a national assessment, further study of the statistical and measurement issues was indicated for 1996.

B. The Puerto Rico Special Assessment

In 1994, a project involving a special adaptation in Spanish of NAEP was conducted. This was the first time a Spanish-language version of the national assessment was ever administered. The findings from the special assessment project in Puerto Rico were of value for NAEP. Because Spanish is the dominant language in Puerto Rico, technically, the students living there are not an LEP population (in fact, the English proficiency of many Puerto Rican students is limited). However, for NAEP, the process of translating and adapting an assessment normally conducted in English, and the careful examination of item characteristics and other data from the assessment, were useful in giving an indication of the procedural implications, psychometric considerations, and the possible problems that may be encountered when doing this type of large-scale assessment.

Factors Leading to a Special NAEP-like Assessment in Puerto Rico

In the early 1990s, education officials in Puerto Rico requested to be part of the next NAEP mathematics and science assessments planned by NCES. Puerto Rico was interested in participating in the Trial State Assessment Program, unfortunately, this was not possible because no versions of the NAEP assessments existed at that time in the Spanish language. In 1993, ETS was asked about the possibility of participating in a special project to conduct NAEP-like assessments in Puerto Rico. After many discussions, ETS contracted with the University of Puerto Rico to conduct assessment in mathematics and science at grades four and eight in 1994. This assessment was called the Puerto Rico Assessment of Educational Progress (PRAEP).

PRAEP was included as a part of a larger project, the PR-Statewide Systemic Initiative, funded under a grant from the National Science Foundation to reform science and mathematics education in Puerto Rico. PRAEP was carried out as a collaborative effort among the Puerto Rico Department of Education, the General Council on Education, and the University of Puerto Rico. The information gathered by PRAEP was used to provide detailed data on what PR's students in grades four and eight knew and could do in mathematics and science in 1994, and to serve as a baseline of data on students' progress in these subject areas.

Design and Operational Issues for the Puerto Rico Assessment of Educational Progress

A hybrid approach of the methods used in the national and TSA assessments was developed for use in PRAEP. Sample size requirements were similar to those used in the 1992 TSA; analysis and reporting procedures were similar also. But, instead of school staff administering the sessions, PRAEP hired and trained its own staff to conduct the assessments. To minimize the effect of language differences, the assessment items and background questions were translated by staff in Puerto Rico into the version of Spanish familiar on the island. Blocks of cognitive items that were used in previous NAEP assessments were adapted for use in PRAEP. Assessment booklets were prepared, similar in appearance in all aspects except for language, to those used in NAEP. In addition, the

complete set of student, teacher, and school background questionnaires were translated and adapted for use in PRAEP. All procedures for administering the assessment and collecting the data were similar to those used in NAEP. The assessment was administered in early 1994.

Following the data collection for the assessment, readers from the island trained by ETS staff scored the mathematics and science constructed-response items in Puerto Rico. Each item was scored using scoring guides similar to those used with the NAEP assessments, but translated into Spanish. Extensive analyses of the data were then conducted, which included examinations of frequency distributions for background questions, item analyses, and differential item functioning analysis for the cognitive items. IRT analyses were performed and the results were placed on a NAEP-like scale for reporting. Two reports were written, one for each of the subject areas (Olson 1995a, 1995b). In addition, a Technical Report was prepared that summarized the operational and technical aspects of this project (Anderson and Olson 1996).

The implementation of the PRAEP project provided valuable information on a number of issues concerning the viability of developing and administering a NAEP-like assessment in another language. In general, researchers found that the activities associated with the implementation of the large-scale assessment project, including the sampling and test administration procedures, were straightforward and could be done without any major problems. In addition, the adaptation of NAEP items into Spanish was mostly successful, so that reliable and valid measures of Puerto Rican students' achievement in mathematics and science could be obtained. However, some problematic areas were identified during the initial field testing of the instruments and procedures (such as errors in translations), and, although these were corrected prior to the actual administration in 1994, other problems were highlighted during the extensive analysis of the assessment data.

Findings from the Puerto Rico Assessment

In the analysis of the 1994 PRAEP data, two areas were of particular concern—the functioning of the Spanish versions of the NAEP items (in comparison to how the items functioned in their original English versions) and the scaling of the results. Results from the *1994 PRAEP Technical Report* described three aspects of the analyses of the PRAEP data that differed from the analyses of the 1990 or 1992 NAEP data (Anderson and Olson 1996).

(1) Few or none of the Puerto Rico student responses fell into the top categories for any of the extended constructed-response items. Therefore, the top two score categories were collapsed into one category.

(2) Because of the difficulty of the items in general for the Puerto Rican students, many of the items did not discriminate well between students at different ability levels, which caused unstable item parameter estimates. Therefore, some of the items had to be jointly calibrated using data from both the Puerto Rico assessment and the NAEP national assessments.

(3) Some items in PRAEP were found to have inappropriate translations or content that where not meaningful for students in Puerto Rico. These items had to be dropped from the PRAEP scales used for reporting.

A series of analyses were conducted to examine the PRAEP data. In the Puerto Rico assessment, conventional analysis of the items took on more importance. Item and test analyses were used to identify items with translation problems, items that did not match the curriculum of Puerto Rico, and items that would pose special problems for the Item Response Theory models used to scale NAEP results. In PRAEP and NAEP, different blocks of items are administered to different students to reduce the response burden, but the blocks are intended to be parallel in what they measure. The item analysis results, summarized by block, indicated that the blocks differed somewhat in average difficulty, reliability, and proportion of students attempting the last item, and thus, the blocks did not appear to be parallel to one another.

A procedure called "Differential Item Functioning (DIF) analysis" was conducted to identify items for which the Puerto Rico and national samples of students (i.e., for the Spanish and English versions) performed differently, controlling for the students' ability levels. For items so identified, wording and context were re-examined with respect to fairness, translation issues, and appropriateness of content before final decisions were made regarding their inclusion in the scales. Results from the DIF analyses led to several items being dropped from scoring and scaling in mathematics and science. Most of these items were deleted due to language issues, such as confusing terminology or different interpretations of the words used in Spanish versus their intended meaning in English.

Based on all of these findings, along with the differences noted between the performance of the Puerto Rico and national samples and a variety of other considerations, the researchers decided that it was not possible to express the IRT results for PRAEP on the same scales as the NAEP results. In other words, the scales established for reporting Puerto Rico results were unique for that jurisdiction. As stated in the chapter on data analysis and scaling from the PRAEP Technical Report, "in order to solve the linear indeterminacy of the PRAEP IRT scales, common-item linking methods were used in establishing the reporting metrics of the PRAEP scales in a way that connected them loosely to the NAEP scales. For this reason, the PRAEP scales permit only limited comparison to NAEP results. Comparison of results for subgroups of students within the 1994 Puerto Rico assessment, however, are completely valid and useful" (Anderson and Olson 1996, p. 63).

C. 1996 NAEP Assessment

For the 1996 assessment of mathematics and science, the U.S. Department of Education and NCES were committed to increasing the participation in the assessment of students with disabilities and those whose English proficiency was limited. New criteria emphasizing inclusion were tried successfully in the 1995 field test and the use of various

accommodations have been proven feasible. These results supported the implementation of these changes in the full 1996 assessment.

Issues remained, however, that complicated the effort to promote inclusion. One of the main reasons for conducting NAEP is to measure progress in academic achievement from assessment to assessment. To measure these trends in a valid way, the assessment had to be administered under the same conditions as in previous assessments. The population of students taking the assessment also had to be selected in the same way. Therefore, the assessment had to be administered both without changes to measure trends, and with the changes needed to promote inclusion.

Further, while the new criteria and accommodations were tried out on a small scale in the field test, no one knew how they would work in a large-scale, full assessment. The 1996 NAEP had to be conducted in such a way that the impact of the new procedures could be measured.

Special Sample Design and Implementation of New Inclusion Criteria and Accommodations

To resolve these problems, the national NAEP sample was divided into three approximately equal subsamples. Each subsample consisted of about 3,500 students each in grades four, eight, and twelve. The subsamples differed in the criteria by which students with disabilities and limited English proficient students were selected for participation in the assessment, and by whether accommodations were offered to these students if needed. This design is illustrated in table 4.3.

Table 4.3—Sample Design for 1996 NAEP Mathematics Assessment

S1: 1992–94 Inclusion Criteria No Accommodations	S2: 1996 Inclusion Criteria No Accommodations	S3: 1996 Inclusion Criteria With Accommodations
Non SD/LEP Students	Non SD/LEP Students	Non SD/LEP Students
Included SD/LEP students (tested without accommodations)	Included SD/LEP students (tested without accommodations)	Included SD/LEP students (tested with or without accommodations)
Excluded SD/LEP students	Excluded SD/LEP students	Excluded SD/LEP students

NOTE: The term SD is used in this table for Students with Disabilities

One of the subsamples, designated S1 in the table, was assessed under the same conditions as in 1994. The inclusion criteria were unchanged, and no accommodations were offered. This subsample was augmented by adding the students without disabilities and non-LEP students in S2, and this combined grouping constituted the "reporting sample" for reporting the main 1996 NAEP results. The results for this reporting sample can be compared to results from past assessments of mathematics in 1990 and 1992, allowing measurement of trends. (For science, subsample S1 was not needed because the science framework was new in 1996. Thus, the science assessment was not comparable to previous science assessments.)

Subsample 2 (S2) was assessed using the new inclusion criteria. Schools were instructed to include students with disabilities unless their cognitive functioning was so severe that they could not be assessed without accommodation or unless their IEP specifically excluded them from testing. Whether they participated in "mainstream" classes during the school year was not to be considered. Students whose English proficiency was limited were to participate in the NAEP if they had completed three years of academic instruction in English, or otherwise, if the school staff thought they could participate in English.

In the third subsample (S3), students with disabilities and limited English proficient students were selected for participation using the new criteria (as with S2), and they were offered accommodations if needed. Accommodations were offered only at the national level, not at the state level.

Accommodations for students with disabilities, in both mathematics and science, were those tested in the 1995 field test, as described above. Students selected in this subsample (S3) were offered the accommodations they normally received as part of their IEP during the school year. For the hands-on portions of the science assessment, physically or visually handicapped students were excluded if they could not carry out the activities.

Limited English proficient students in subsample S3 were given a bilingual Spanish-English booklet in the mathematics assessment if needed. Based on the findings from the field test, a Spanish-only booklet was not offered. Because no science translation of the items had been field-tested in 1995, LEP students were not offered a bilingual booklet as part of the science assessment. However, they were offered an English-Spanish glossary of scientific terms.

The procedures described thus far are pertinent to the national-level NAEP assessment. NAEP is also administered at the state level for states that voluntarily participate. In 1996, 47 jurisdictions (including the District of Columbia and Guam) participated at grade four and 44 jurisdictions participated at grade eight. In these jurisdictions, mathematics was assessed at grades four and eight, and science was assessed at grade eight. While the national-level NAEP is administered and funded by the federal government, each state administers and funds its own state-level assessment. Because agreements had already been completed with the states, who had volunteered earlier to participate in a 1996 assessment

that did not include accommodations, it was decided not to impose the changes that would be required for the extra administrative and financial burden of providing accommodations at the state level. Therefore, to enable evaluation of the impact of changes over time at the state level for the revised inclusion criteria, a split-sample approach was used with the pre-1996 criteria applied in half the schools, and the new 1996 criteria applied in the other half.

Reporting 1996 Results

Comparing average results and inclusion rates among the three subsamples would measure the effects on NAEP overall of using the new inclusion criteria and providing accommodations. At the same time, results from previous years in mathematics can be compared with results from the 1996 reporting sample (defined above). It was hoped that implementation of the new procedures would result in increased participation of students with disabilities and limited English proficient students. Initial results pertaining to these students, reported in the *NAEP 1996 Mathematics Report Card for the Nation and the States* (February 1997) and in the *NAEP 1996 Science Report Card for the Nation and the States* (May 1997), indicate that participation rates did increase appreciably when accommodations were offered, although the changed inclusion criteria did not seem to make a difference in participation rates. The NAEP Mathematics and Science Report Cards summarize the overall national and state results, reporting average scores of all students and selected subgroups of students. Additionally, the NAEP Report Cards have a separate chapter describing the procedures used to include more students with special needs in NAEP. The reports also present participation rates of students with disabilities and limited English proficient students at the national and state levels.

Within a year following the release of the 1996 NAEP Report Cards, a separate report is planned to make available findings on the effects of the efforts to increase the participation of students with disabilities and limited English proficient students in the assessments. This report will present data on the effects on overall scale score averages of the new participation guidelines and the provision of accommodations. Separate average scores for students with disabilities and limited English proficient students will not be provided, however.

This special report will also include studies of how well the results for accommodated students' fit on the NAEP scale, and how the patterns of responses given by these students compare with those of non-accommodated students. There also will be information on the frequency of the various types of disabilities encountered, the number of students accommodated in the sample, and the frequency with which the various accommodations were used. In addition, the report will describe background characteristics of students with disabilities and limited English proficient students.

D. Summary

As discussed in this chapter, NCES is in the process of examining and implementing new procedures for NAEP to maximize the representation of students included in the assessment. The focus is on efforts to enhance overall participation by including more students with disabilities and LEP students. Changes made in the criteria for making selection decisions will hopefully result in more clarity in the judgments made for including students in the assessment and in more uniformity across jurisdictions. Higher participation rates have also been encouraged by offering various accommodations to students with special needs.

The efforts to increase the participation of students with disabilities and limited English proficient students in NAEP have been an evolving process. In the Puerto Rico Assessment of Educational Progress, conducted in 1994, techniques and procedures were developed leading to provision of accommodations for limited English proficient students in the 1996 NAEP. The 1995 NAEP field test served to refine these accommodations, and to introduce new procedures and inclusion criteria to facilitate the participation of students with disabilities. Finally, the 1996 NAEP incorporated a research component, enabling the measurement of the effects of new inclusion criteria and accommodations on overall results, as well as maintaining the ability of NAEP to accurately measure trends in student achievement.

In the final chapter of this report, ongoing research into the issues raised by the changes in design and procedures needed to enhance inclusion in large-scale assessments such as NAEP are described, and additional technical issues that still need to be resolved are discussed.

Chapter

5

Ongoing Studies, Research Needs, and Future Directions

A. NCES's ongoing research studies on inclusion issues
   - Scaling issues
   - Reporting issues
   - Construct validity issues
   - Language complexity issues

B. Further technical and measurement issues
   - General measurement questions and technical issues for large-scale assessments
   - Appropriateness of inclusion criteria and procedures
   - Impact of accommodations on assessment results
   - Reporting of disaggregated data

C. Overview of the next volume of this report

D. Conclusion

A. NCES's Ongoing Research Studies on Inclusion Issues

In chapter 4, the 1996 NAEP sampling plan was described as a large-scale research design. The use of three subsamples, differing in testing conditions, made it possible to measure the impact of changes in criteria for inclusion of students with disabilities and LEP students, and the impact of accommodations introduced to remove barriers to the participation of these students in the assessment. In this section, descriptions of the ongoing research projects to study the effects of implementing the inclusion strategies in NAEP are presented.

Because this is a pioneering effort by NAEP, and examination of the topic is difficult because of its many complexities, additional research is required. A number of studies, funded in part by contributions from OERI, OSERS and OBEMLA, are underway to investigate further inclusion issues. Some of the studies are incorporated into the analysis of NAEP data by NAEP's main contractor, Educational Testing Service. Various studies also are being implemented by other researchers, independent of ETS. This collection of studies

is intended to further NAEP's progress toward the goal of full inclusion, while preserving the overall validity of the assessment. Research studies like these are essential to ensure that the reporting of what students know and can do, as well as gauging their academic progress over the years, is done in a reliable, valid, and meaningful way. As the various issues are examined and new procedures implemented, NAEP must also maintain its commitment to high technical standards, the continuity of its trend data, and a balance of resources for the program.

The collection of inclusion research studies based on the 1996 NAEP are grouped, as follows, according to the main topic and issues under investigation:

- *Scaling—does the NAEP scale accurately reflect the results for respondents assessed under non-standard conditions?*

  As the test booklets from students assessed under standard and nonstandard (i.e., accommodated) conditions are scored, the patterns of responses of these groups are plotted onto graphs. An item characteristic curve (ICC) is plotted for each item and shows the ability levels of students by the probability of answering the item correctly. In the usual pattern shown on the ICCs, as ability level rises, the probability of getting any particular item correct also increases. During ETS's analysis of the 1995 field test of NAEP, it was found that the ICCs of many items did not follow the usual pattern when plotted for accommodated students. For these items, the probability of getting the correct answer did not rise with ability level, or it rose at a different rate than for nonaccommodated students, or it followed a different curve entirely. This may indicate that the results for the accommodated students do not mean the same as the results for nonaccommodated students. The reasons for the differences are not fully known, but this finding prevented the data for accommodated students from being included with the data for other students on the NAEP scale.

  ETS will subject the results for accommodated students in the 1996 assessments to the same type of analysis, although certain features of the 1996 assessment should permit more definitive research findings than those from the 1995 field test. The full assessment drew upon a larger sample of accommodated students than the field test. In addition, all accommodated students were given the same test booklet, in order to yield a reasonably large sample of respondents for each of the items in that booklet. It is possible that the larger sample will lead to response patterns for accommodated students that are closer to those for students assessed under standard conditions.

- *Reporting—if results for accommodated students cannot be reported on the NAEP scale, how might these results best be reported?*

  Because of the possibility that it may be difficult to scale the results for accommodated students, ETS is examining alternatives to reporting results on the 500-point NAEP scale. One such alternative is simply to report the average percent of students who answered a particular item correctly. In the case of constructed-response

items, for which the student must give a written answer rather than choose a correct response from among given options (multiple choice), results could be presented in terms of the percent of students who gave partially correct and complete answers. Advantages of this approach are that these results would be easily understood and would offer a way of reporting for students whose results otherwise could not be presented, and results could be provided for individual items or possibly for collections of items. Disadvantages of this approach are that the percent-correct statistics could not be reported on the NAEP scale and would not provide as much information as scaled results, nor would they provide as stable and comparable measures as scaling does.

- *Construct validity—do NAEP results mean the same for accommodated students as for nonaccommodated students?*

  The findings from the 1995 field test did not clearly indicate if the results for accommodated students will fit on the NAEP scale. An answer to this question is important because of concerns that the assessments may not be measuring the same constructs for these students as for students assessed under standard conditions, which would mean that inferences made from the results for accommodated students about their knowledge and abilities may not be valid.

  Studies are currently in progress to determine whether the accommodations offered to LEP students are in some way distorting the validity of the assessment. As discussed in chapter 3, researchers at CRESST are investigating the relative impact of linguistic alternatives to the wording of NAEP items for both LEP and non-LEP students. They are examining the use of simplified English wording, translation to Spanish, and other aids to understanding, such as pictures and oral reading and answering. The researchers believe that these accommodations should not result in any substantial change to the results achieved by non-LEP students. If they do, it would imply that these accommodations may be affecting the validity of the measures of the assessment. At the same time, CRESST will examine which accommodations lead to improved measurement of LEP students' knowledge of the subject matter (mathematics), separate from their ability to read and answer in English.

  In addition, as noted in the discussion of the proposed NAEP Validity Studies research project in chapter 2, a study of the validity of accommodations for students with disabilities is being designed. This study would examine the effect of accommodations, extended time in particular (the most often-used accommodation), on the response patterns of students with disabilities and students without disabilities. If the accommodations led to different results for students without disabilities (compared to taking the assessment under standard, i.e., non-accommodated, conditions) the validity of using the accommodations would need to be re-examined.

- *Language complexity—does the proficiency of LEP students in their native language affect their performance on NAEP, and do the linguistic patterns of written English responses to constructed-response items of LEP students lead to distorted scoring?*

  In addition to the study discussed above, CRESST is also studying the impact of native language proficiency on the performance of LEP students in NAEP. The project, conducted with Spanish-speaking LEP students, will examine whether their degree of literacy in their native language has an impact on their performance on NAEP mathematics items in either English or Spanish.

  As discussed in chapter 3, the LEP SCASS group of the CCSSO is designing a study to evaluate the impact of training item scorers to be sensitive to psycho-linguistic patterns of responses in English given by LEP students to extended-constructed response items (i.e., items that require more lengthy written answers). It may be that the content knowledge of some LEP students is masked by their lack of English-language mastery, resulting in correct written answers wrongly judged as incorrect by scorers. Misspelling and distorted syntax may make these students' answers unintelligible to scorers not trained to recognize these patterns. The research team intends to train a group of scorers, using a special manual designed to overcome this problem. They would determine whether this supplemental training yields more accurate scoring of extended-response items for LEP students.

B. Further Technical and Measurement Issues

In addition to the topics discussed in this report, a number of other issues, technical and otherwise, need to be addressed further in order to gain a better understanding of the implications of providing accommodations. Although the findings from recent and ongoing studies are helping to answer some of the questions that have been addressed, many challenges still exist that may stand in the way of best measurement practice and the proper implementation of assessment methodologies that are technically sound.

NAEP alone cannot answer all of the questions that exist. Different types of large-scale assessments are in use in many different localities, some with very different approaches and purposes than NAEP. Because there are limits to the answers that can be found from the ongoing collection of studies, more research is needed at the national, state, and local levels to further examine some of the following issues.

General Measurement Questions and Technical Issues for Large-Scale Assessments

Measurement issues constitute one of the biggest challenges when implementing a more inclusive assessment. Obviously, assessments need to meet certain standards to provide accurate measures of student achievement or performance (AERA/APA/NCME 1985). Some of the key questions that need to be addressed when considering the technical adequacy of any large-scale assessment are:

- Is the assessment of high technical quality? What are the psychometric characteristics of the assessment (e.g., level of performance, test completion rates, internal-consistency reliability, differential item performance) for students with disabilities and LEP students, as well as for all students?

- Does the assessment provide students with disabilities and LEP students a fair opportunity to answer questions across the range of difficulties being tested? In particular, for lower performing students, does the test have an adequate number of items to measure ability at the low end of the scale?

- Does the assessment provide a reliable and consistent measure of students with disabilities and LEP student performance? Is the reliability of an accommodated assessment the same as that for the assessment given under standard conditions?

- Is the construct validity of the assessment the same for students with disabilities or LEP students and other students? Does the assessment measure the same constructs for the different groups? Is the validity of an accommodated assessment the same as that for the assessment given under standard conditions?

- Are the scores obtained by students with disabilities or LEP students comparable to those of other students? Do the scores for the different groups mean the same thing? Is the factor structure the same for the different groups? Can scores for the different groups be put on the same scale for reporting?

Some of the other technical issues that may need to be examined further pertain to the predictive validity of the assessment (if predictions are appropriate) or to admissions or pass/fail decisions made based on the results. Although many of these issues were studied in the ETS book *Testing Handicapped People* (Willingham et al. 1988), the focus was on students with disabilities and not LEP students, and the findings pertained more to college and graduate school admissions tests. Since 1988, the use of large-scale assessments at the local and national levels has grown dramatically. In addition, the current wave of assessments are increasingly molded to do more things and to include more people. Thus, many of the questions that were studied in the previous decade need to be revisited and answers found that are relevant to the latest generation of assessment approaches.

Appropriateness of Inclusion Criteria and Procedures

Additional research may be needed to further determine the appropriateness of the inclusion criteria and procedures used for students with disabilities and LEP students in large-scale assessments. Studies, for example, that examine the proper role of the IEP team for students with disabilities, or whether the threshold for including English-language learners in large-scale English language assessments should be two or three years for LEP students, may help resolve some of the existing questions. In addition, an investigation to determine if there still are students being excluded who could be included in the assessment, and how the different

inclusion criteria are being applied by local staff, could provide valuable information on the efficacy of the inclusion procedures.

Impact of Accommodations on Assessment Results

Another area that may need to receive additional attention is how accommodations may affect student performance in large-scale assessments. As discussed in previous chapters, a wide variety of assessment accommodations are being used. Much data are being collected on the extent of the use of test accommodations at the local and national levels; the impact of their use on assessment measures, however, is not always known in all situations. In particular, the impact of additional time to take the test is in question. Additional research is needed to determine the effect of extra time on test performance, for both students with disabilities or LEP students as well as for other students.

For LEP students, another area receiving attention is adapting or developing tests in students' home languages. Of the accommodations typically used to include more LEP students, alternate language assessment, increased testing time, and provision of glossaries to aid students whose native language is not the test language are most common. However, implementing these provisions pose some formidable challenges. Unfortunately, some of the existing assessments in other languages are not always of high quality. Nor are they necessarily comparable in scoring or difficulty or even content to their companion tests in English. Also, assessments are not yet available in most languages. If assessments in different languages are developed, they must meet technical standards and be comparable to the assessments being conducted for English-speaking students. The results of research on the technical issues associated with alternate-language assessments are mixed, and no consensus exists on the most appropriate direction to proceed, including whether to translate assessments at all.

Reporting of Disaggregated Data

Another challenge that needs to be met concerns how data for students with disabilities and LEP students should be reported. For example, some state assessment programs are designed and used for purposes of school and district accountability, and these programs may have specified policies for the use of assessment results that require the reporting of disaggregated data. However, questions still exist on how data should be disaggregated and reported, at the local, state, and national level. Sample size requirements for reporting of subgroup performance and the accuracy of the measures being reported are key issues that must be taken into account. In addition, concerns exist, especially in small districts and schools, of the possibility that the disaggregation of results could cause breaches of privacy, for example, where there are very few students who fall into a specific classification that is reported.

C. Overview of the Next Volume of this Report

Currently, it may seem that there are more questions than answers on the issues associated with the inclusion of students with disabilities and LEP students in large-scale assessments. At this point in time, many different studies are underway and results are forthcoming (some, but not all of which, are being addressed in the currently ongoing NAEP studies). The research projects supported by NCES and other areas of the U.S. Department of Education will help answer some of the questions. It is expected that other studies, being done independently by various states and by other organizations, will also provide valuable information.

This report is the first part of a series that focuses on the projects, activities, and findings on the inclusion of students with disabilities and LEP students in large-scale assessments. In this volume, an overview of the recent history leading to the increased focus on more inclusive assessments was presented and many of the studies recently completed and currently underway were summarized. The focus in this volume has been intentionally broad, with summaries of many projects being conducted at the local and state levels, as well as at the national level. The second volume, planned for release in late 1997, will summarize the results from a number of these studies that focus specifically on NAEP data. These studies include the work that is being done to examine the possibility of scaling the results from accommodated students, determining the impact of revising the inclusion criteria used in the assessment on NAEP's ability to measure trends, examining the effect of providing accommodations on the measures obtained by NAEP, and evaluating the comparability of results for students who received accommodations with those who did not. The second volume will conclude by providing suggestions for further research on some of the unresolved issues. These recommendations are intended to help guide NCES in its overall approach toward further developing more inclusive large-scale assessments and surveys. Although the focus of the next volume will be primarily on NAEP, the findings presented will be of value to many other assessment and survey programs.

D. Conclusion

The changes incorporated into the 1996 NAEP that further the goal of maximum inclusion of students with disabilities and LEP students will result not only in an improved national assessment program, but will also benefit states, school districts, and other entities that conduct large-scale assessments. Many educators at these levels look to NAEP as a model for the best practice in assessment. Thus, NAEP needs to proceed in a thoughtful and thorough manner in its implementation of a more inclusive assessment.

Issues of fairness, equity, representativeness, and accuracy are of utmost concern to NCES as it implements new procedures for NAEP. NCES is committed to increasing the accessability of the assessment in order to enhance the overall representativeness and generalizability of the findings. A main goal is to provide good data on student achievement, data which include all students in the measures. To meet this goal, assessments need to be

designed that enable student performance in ways that are appropriate for everybody. In addition, as required, accommodations need to be provided that help to level the playing field. NCES is making progress toward these goals. As the procedures and psychometric characteristics of a more inclusive assessment continue to be closely examined, attainment of these goals will have to be balanced with the challenges of maintaining a valid and reliable assessment. This is the challenge for all who are involved in this topic.

REFERENCES

Bibliography

Abedi, J. et al. *Language Background as a Variable in NAEP Mathematics Performance: NAEP TRP Task 3d—Language Background Study*. University of California, National Center for Research on Evaluation, Standards, and Student Testing, Technical Review Panel for Assessing the Validity of the National Assessment of Educational Progress, May 1995.

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. *Standards for Educational and Psychological Testing*. Washington, DC: American Psychological Association, 1985.

American Psychological Association Joint Committee on Testing Practices. *Assessing Individuals with Disabilities*. Washington, DC: American Psychological Association, forthcoming.

Anderson, N.E., Jenkins, F.F., and Miller, K.E. *NAEP Inclusion Criteria and Testing Accommodations: Findings from the NAEP 1995 Field Test in Mathematics*. Princeton, NJ: Educational Testing Service, 1996.

Anderson, N. and Olson, J. *Puerto Rico Assessment of Educational Progress: 1994 PRAEP Technical Report*. Princeton, NJ: Educational Testing Service, August 1996.

August, D. and Hakuta, K. *Improving Schooling for Language-Minority Children: A Research Agenda*. National Research Council, National Academy Press, Washington, DC, 1997.

August, D., Hakuta, K., and Pompa, D. "For All Students: Limited English Proficient Students and Goals 2000." *Occasional Papers in Bilingual Education*, 10 (1994): 4.

Bennet, R.E. "Testing Examinees with Handicaps for College and Graduate Admissions." *Special Services in Schools*, 2 (1984): 21-29.

——*Computer-Based Testing for Examinees with Disabilities: On the Road to Generalized Accommodation*. ETS Research Memorandum No. RM95-1. Princeton, NJ: Educational Testing Service, May 1995.

Bennet, R.E. and Ragosta, M. "The Technical Characteristics of Postsecondary Admissions Tests for Handicapped Examinees: A Review of Research." *Journal of Special Education*, 19 (1985): 255-267.

Bennet, R.E., Rock, D.A., and Ignatkoski, I. "Differential Item Functioning on the SAT-M Braille Edition." *Journal of Educational Measurement*, 26 (1989): 67-79.

Bennet, R.E., Rock, D.A., and Jirele, T. "GRE Score Level, Test Completion, and Reliability for Visually Impaired, Physically Handicapped, and Nonhandicapped Groups." *Journal of Special Education,* 21 (1987): 9-21.

——"Sex Group Membership as a Confounding Factor in Handicapped Students GRE General Test Performance." *Journal of Postsecondary Education and Disability,* 7 (1989): 61-68.

Bennet, R.E., Rock, D.A., and Kaplan, B. A. "Level, Reliability, and Speededness of SAT Scores for Nine Handicapped Groups." *Special Services in Schools,* 3/4 (1988): 37-54.

——"SAT Differential Item Performance for Nine Handicapped Groups." *Journal of Educational Measurement,* 24 (1987): 41-55.

Bond, L.A., Council of Chief State School Officers and North Central Regional Educational Laboratory. "Statewide Assessment of Students with Disabilities." Oak Brook, IL: North Central Regional Educational Laboratory, January 1996.

Bond, L.A., Braskamp, D., and Roeber, E. *The Status Report of Assessment Programs in the United States: State Student Assessment Programs Database, School Year 1994-1995.* Washington, DC: North Central Regional Educational Laboratory and Council of Chief State School Officers, 1996.

Bradby, D. *Language Characteristics and Academic Achievement: A Look at Asian and Hispanic Eighth Graders in NELS: 88.* Washington, DC: National Center for Education Statistics, 1992.

Braun, H., Ragosta, M., and Kaplan, B. *The Predictive Validity of the Scholastic Aptitude Test for Disabled Students.* ETS Research Report No. 86-42. Princeton, NJ: Educational Testing Service, 1986a.

——*The Predictive Validity of the GRE General Test for Disabled Students.* ETS Research Report No. 86-38. Princeton, NJ: Educational Testing Service, 1986b.

Cheung, O., Clements, B., and Miu, Y.C. *The Feasibility of Collecting Comparable National Statistics about Students with LEP.* Washington, DC: National Center for Education Statistics, 1994.

Committee on Goals 2000 and the Inclusion of Students with Disabilities. *Annual Report.* Washington, DC: National Academy of Sciences, National Research Council, Commission on Behavioral and Social Sciences and Education, Board on Testing and Assessment, October 1995.

Council of Chief State School Officers. *1996-97 SCASS and CIIP Project Descriptions: State Collaborative on Assessment and Student Standards Curriculum and Instructional*

Improvement Program. Washington, DC: Council of Chief State School Officers, May 1996.

Council of Chief State School Officers and North Central Regional Educational Laboratory. *1996 State Student Assessment Programs Database.* Oak Brook, IL: North Central Regional Educational Laboratory, 1996.

Educational Testing Service, ETS Committee on People with Disabilities. *Testing Persons with Disabilities: A Report for ETS Programs and their Constituents.* Princeton, NJ, Educational Testing Service, undated.

Fleischman, H.L. and Hopstock, P.J. *Descriptive Study of Services to Limited English Proficient Students, Volume 1 Summary of Findings and Conclusions.* Arlington, VA: Development Associates, Inc., 1993.

George Washington University Center for Equity and Excellence in Education. *Promoting Excellence: Ensuring Academic Success for Limited English Proficient Students.* Arlington, VA: The George Washington University Center for Equity and Excellence in Education, Evaluation Assistance Center East, 1996.

Hakuta, K. and Valdes, G. *A Study Design to Evaluate Strategies for the Inclusion of LEP Students in the NAEP State Trial Assessment.* Washington, DC: National Center for Education Statistics, Office of Educational Research and Improvement, U.S. Department of Education, 1994.

Hopstock, P.J. and Bucaro, B.J. *A Review and Analysis of Estimates of the LEP Student Population*, Arlington, VA: Development Associates, Special Issues Analysis Center, 1993.

Houser, J. *Assessing Students with Disabilities and Limited English Proficiency.* Working Paper Number 95-13. Washington, DC: National Center for Education Statistics, Policy and Review Branch, Data Development Division, March 1995.

Jones, D.H. and Ragosta, M. *Predictive Validity of the SAT on Two Handicapped Groups: The Deaf and the Learning Disabled.* Research Report 82-9. Princeton, NJ: Educational Testing Service, 1982.

LaCelle-Peterson, M. & Rivera, C. "It is real for all kids? A framework for equitable assessment policies for English language learners" *Harvard Educational Review, 64(1): 55-75, 1994.*

National Academy of Education. *The Trial State Assessment: Prospects and Realities.* Stanford, CA: National Academy of Education, 1993.

——*Quality and Utility: The 1994 Trial State Assessment in Reading.* Stanford, CA: National Academy of Education, 1996.

National Academy of Sciences/National Research Council: Committee on the Evaluation of National and State Assessments of Educational Progress. *Draft Study Plan*, Washington, DC: National Academy of Sciences/National Research Council, July 1996.

National Center on Educational Outcomes. *State Practices in the Assessment of Outcomes for Students Disabilities.* Report 1. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes, October 1991.

——*Inclusion of Students with Disabilities in National and State Data Collection Systems.* Report 2. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes, March 1992.

——*Experts' Opinions on National Mathematics Standards for Students with Disabilities.* Report 3. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes, September 1992.

——*IEPs and Standards: What They Say for Students with Disabilities*. Report 5. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes, May 1993.

——*The Identification of People with Disabilities in National Data Bases: A Failure to Communicate*. Report 6. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes, May 1993.

——*Secondary Analysis of State Assessment Data: Why We Can't Say Much About Students with Disabilities*. Report 10. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. August, 1994.

——*Disability Summary Analyses of Select National Data Collection Programs*. Report 11. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes, August 1995.

——*Why We Can't Say Much About the Status of Students with Disabilities During Educational Reform.* Report 21. Minneapolis, MN, National Center on Educational Outcomes, August 1995.

——*A Disability Perspective on Five Years of Education Reform.* Report 22. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes, December 1995.

——*Testing Accommodations for Students with Disabilities: A Review of the Literature*. Synthesis Report 4. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes, March 1993.

——*Can "All" Ever Really Mean "All" in Defining and Assessing Student Outcomes.* Synthesis Report 5. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes, March 1993.

——*Implications of Outcomes-Based Education for Children with Disabilities.* Synthesis Report 6. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes, April 1993.

——*Views on Inclusion and Testing Accommodations of Students with Disabilities.* Synthesis Report 7. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes, September 1993.

——*The Effects of Standards and Assessment on Students in Special Education.* Synthesis Report 10. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes, October 1993.

——*National Goals, National Standards, National Test: Concerns for All (Not Virtually All) Students with Disabilities.* Synthesis Report 11. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes, March 1994.

——*Implementation of Alternative Methods for Making Educational Accountability Decisions for Students with Disabilities.* Synthesis Report 12. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes, March 1994.

——*Making Decisions About the Inclusion of Students with Disabilities In Large-Scale Assessments.* Synthesis Report 13. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes, April 1994.

——*Recommendations for Making Decisions About the Participation of Students with Disabilities in Statewide Assessment Programs.* Synthesis Report 15. Minneapolis, MN: National Center on Educational Outcomes, 1994.

——*Compilation of State's Guidelines for Including Students with Disabilities in Assessments.* Synthesis Report 17. Minneapolis, MN: National Center on Educational Outcomes, 1995.

——*Compilation of State's Guidelines for Accommodations in Assessments for Students with Disabilities.* Synthesis Report 18. Minneapolis, MN: National Center on Educational Outcomes, 1995.

——*A Perspective on Education and Assessment in Other Nations: Where are Students with Disabilities*? Synthesis Report 19. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes, April 1995.

——*State Graduation Requirements and Students with Disabilities*. Synthesis Report 20. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes, May 1995.

——*Neglected Numerators, Drifting Denominators, and Fractured Fractions: Determining the Participation Rates for Students with Disabilities in Statewide Assessment Programs*. Synthesis Report 23. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes, 1996.

——*Guidelines for Inclusion of Students with Disabilities in Large-Scale Assessments*. Policy Directions No. 1. Minneapolis, MN: National Center on Educational Outcomes, 1994.

——*State Special Education Outcomes 1994*. Minneapolis, MN: National Center on Educational Outcomes, 1995.

——"The Exclusion of Students with Disabilities in National Data Collection Programs." *Educational Evaluation and Policy Analysis, 15: 339-352*. Minneapolis, MN: National Center on Educational Outcomes, 1992.

——*The Inclusion of Students with Disabilities in National Data Collections Programs*. Technical Report 2. Minneapolis, MN: National Center on Educational Outcomes, 1992.

——*Tough Questions About Accountability Systems and Students with Disabilities in Large-Scale Assessments: Characteristics and Considerations*. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes, 1996.

——*Project Directory. Assessment Projects and Related Efforts*. Minneapolis, MN: University of Minnesota, College of Education and Human Development, National Center on Educational Outcomes, August 1996.

National Clearinghouse for Bilingual Education. *Summary Report of the Survey of the States' Limited English Proficient Students and Available Educational Programs and Services 1994-1995*. Washington, DC: George Washington University, 1996.

North Central Regional Educational Laboratory. *State Student Assessment Programs Database 1994-1995*. Oak Brook, IL: National Central Regional Educational Laboratory, 1995.

Olson, J.F. *Puerto Rico Assessment of Educational Progress: Final Report, 1994 Mathematics Assessment*. Princeton, NJ: Educational Testing Service, July 1995.

——*Puerto Rico Assessment of Educational Progress: Final Report, 1994 Science Assessment*. Princeton, NJ: Educational Testing Service, August 1995.

——(a proposal in conjunction with the NAEP Validity Studies Panel). *A Proposed Study Design to Examine the Inclusion of Students with Disabilities in NAEP: Impact of*

Increased Testing Time on the Performance of Students With and Without Disabilities (Draft). Washington, DC: NAEP Validity Studies Panel, January 1997.

Olson, J.F. and A.A. Goldstein. *Focus on NAEP. Increasing the Inclusion of Students with Disabilities and Limited English Proficient Students in NAEP.* NCES 96-894. Washington, DC: National Center for Education Statistics, July 1996.

Packer, J.K. *SAT Testing Time for Students with Disabilities.* ETS Research Report No. 87-37. Princeton, NJ: Educational Testing Service, 1987.

Phillips, G. *"The National Center for Education Statistics: Development Work and Research on Adaptations and Alternative Assessments for Students with Disabilities."* A presentation given at a Council of Chief State School Officers SCASS Special Education Assessment Meeting in Tampa, FL, January 19, 1996.

Phillips, S.E. *Legal Implications of High-Stakes Assessment: What States Should Know.* Regional Information Center Report. Oak Brook, IL: North Central Regional Educational Laboratory, 1994.

——*All Students, Same Test, Same Standards: What the New Title I Legislation Will Mean for the Educational Assessment of Special Education Students.* Oak Brook, IL: North Central Regional Educational Laboratory, 1995.

Ragosta, M. *Students with Disabilities: Four Years of Data from Special Test Administrations of the SAT 1980-83.* College Board Report No. 87-2 and ETS Research Report 87-2. New York: College Entrance Examination Board, 1987.

Ragosta, M. and Kaplan, B. *A Survey of Handicapped Students Taking Special Test Administrations of the SAT and GRE.* ETS Research Report No. 86-5. Princeton, NJ: Educational Testing Service, 1986.

Ragosta, M. and Wendler, C. *Eligibility Issues and Comparable Time Limits for Disabled and Nondisabled SAT Examinees. E*TS Research Report No. 92-35. Princeton, NJ: Educational Testing Service, 1992.

Ragosta, M., Braun, H., and Kaplan, B. *Performance and Persistence: A Validity Study of the SAT for Students with Disabilities.* College Board Report No. 91-3 and ETS Research Report No. 91-41. New York: College Entrance Examination Board, 1991.

Rivera, C. and Vincent, C. *State Policies and Practices for Testing LEP Students for High School Graduation.* Presentation made at the National Association for Bilingual Education meeting, Orlando, FL, 1996.

——"High school graduation testing: Policies and practices in the assessment of limited English proficient students." *Education Assessment* (In Press).

Rivera, C., Vincent, C., Hafner, A., and LaCelle-Peterson, M. *Statewide Assessment Programs: Policies and Practices for the Inclusion of Limited English Proficient Students—Findings from a National Survey of State Assessment Directors.* (TM #026831). Washington, DC: ERIC Clearinghouse on Assessment and Evaluation. The Catholic University of America, 1997.

Rock, D.A. and Chan, K.L. "SAT Verbal-Math Discrepancies: An Accurate Indicator of College Learning Disability?" *Journal of Learning Disabilities,* 20 (1987): 189-192.

Rock, D.A., Bennet, R.E. and Jirele, T. "The Factor Structure of the Graduate Record Examinations General Test in Handicapped and Nonhandicapped Groups." *Journal of Applied Psychology,* 73 (1988): 383-392.

Rock, D.A., Bennet, R.E., and Kaplan, B.A. "Internal Construct Validity of a College Admissions Test Across Handicapped and Nonhandicapped Groups." *Educational and Psychological Measurement,* 47 (1987): 193-205.

Secada, W.G. "Issues in the Development of Spanish-Language Versions of the National Assessment of Educational Progress." In *The Trial State Assessment: Prospects and Realities: Background Studies.* Ed. George Bohrnstedt. Palo Alto, CA: Armadillo Press, 1994.

Shepard, L.A. "Evaluating Test Validity." *Review of Research in Education* 19 (1993): 405-450.

——"Research Framework for Investigating Accommodations for Language Minority Students." Presentation made at CRESST Assessment Conference, UCLA, 1996.

Spencer, B. "A Study of Eligibility Exclusions and Sampling: 1992 Trial State Assessment." In *The Trial State Assessment: Prospects and Realities: Background Studies.* Ed. George Bohrnsted. Palo Alto, CA: Armadillo Press, 1994.

——"School and Student Sampling in the 1994 Trial State Assessment: An Evaluation." In *Quality and Utility: The 1994 Trial State Assessment in Reading, Background Studies.* Stanford, CA: National Academy of Education, 1996.

Stancavage, F., et al. "Study of Exclusion and Assessability of Students with Limited English Proficiency in the 1994 Trial State Assessment of the National Assessment of Educational Progress." In *Quality and Utility: The 1994 Trial State Assessment in Reading, Background Studies.* Stanford, CA: National Academy of Education, 1996.

Stancavage, F., et al. "Study of Exclusion and Assessability of Students with Disabilities in the 1994 Trial State Assessment of the National Assessment of Educational Progress." In

Quality and Utility: The 1994 Trial State Assessment in Reading, Background Studies. Stanford, CA: National Academy of Education, 1996.

U.S. Department of Education. *Tenth Annual Report to Congress on the Implementation of the Education of the Handicapped Act.* Washington, DC: U.S. Department of Education, 1988.

———*PROSPECTS: The Congressionally Mandated Study of Educational Growth and Opportunity, Interim Report: Language Minority and Limited English Proficient Students.* Washington, DC: U.S. Department of Education, 1993.

———*PROSPECTS: The Congressionally Mandated Study of Educational Growth and Opportunity, First Annual Report.* Washington, DC: U.S. Department of Education, 1996.

———*The Condition of Education 1996.* NCES 96-304. Washington, DC: Office of Educational Research and Improvement, U.S. Department of Education, 1996.

———*Research in the Education of Individuals with Disabilities. Innovation and Development in Special Education: Directory of Current Projects 1996 Edition.* Washington, DC: Office of Special Education Programs, Division of Innovation and Development, U.S. Department of Education, 1996.

———*To Assure the Free Appropriate Public Education of All Children with Disabilities.* Eighteenth Annual Report to Congress on the Implementation of the Individuals with Disabilities Act, Washington, DC: Office of Special Education Programs, U.S. Department of Education, 1996.

U.S. Department of Education, National Center for Education Statistics. *National Education Longitudinal Study of 1988, Base Year: School Component Data File User's Manuals.* Washington, DC: Government Printing Office, March 1990.

———*1992 Trial State Assessment: Summary of Participation Rates, Training, and Data Collection Activities.* Washington, DC: National Center for Education Statistics, July 1992.

———*1994 Trial State Assessment: United States Report on Data Collection.* Washington, DC: National Center for Education Statistics, October 1994.

———"Conference on Inclusion Guidelines and Accommodations for Limited English Proficient Students in the National Assessment for Educational Progress." Briefing Book. Washington, DC: December 5, 1994.

———*Report on Data Collection Activities for the 1994 National Assessment of Educational Progress.* Washington, DC: National Center for Education Statistics, March 1995.

———*Proceedings of the Conference on Inclusion Guidelines and Accommodations for Limited Proficient Students in the National Assessment of Educational Progress*, by August,

D. and McArthur, E. NCES 96-861. Washington, DC: Office of Educational Research and Improvement, National Center for Education Statistics, U.S. Department of Education, May 1996.

United States Congress. *Individuals with Disabilities Education Act.* PL 94-142. Washington, DC: Government Printing Office, 1990.

———*The Americans with Disabilities Act of 1990,* PL 103-336, as amended by the Civil Rights Act of 1991. Washington, DC: Government Printing Office, 1991.

———*Improving America's Schools Act.* PL 103-382. Washington, DC: Government Printing Office, 1993.

———*Goals 2000: Educate America Act.* PL 103-227. Washington, DC: Government Printing Office, 1993.

Willingham, W.W. *Handicapped Applicants to College: An Analysis of Admissions Decisions.* College Board Report No. 87-1 and ETS Research Report No. 87-1. New York: College Entrance Examination Board, 1987.

——"Standards Testing Conditions and Standard Score Meaning for Handicapped Examinees." *Applied Measurement in Education,* 2(2) (1989): 97-103.

Willingham, W.W., Bennet, R.E., Braun, H., Powers, D.E, Ragosta, M., and Rock, D.A. *Testing Handicapped People.* Sponsored by the College Board, Educational Testing Service, Graduate Record Examinations Board. Needham Heights, MA: Allan and Bacon, Inc., 1988.

Yesseldyke, J. "Overview of the Issues of Participation and Inclusion of Students with Disabilities in Large-Scale Assessment Programs." A presentation at a Council of Chief State School Officers SCASS Special Education Assessment Meeting in Tampa, FL, January 19, 1996.

Zehler, A.M., et al. *Literature Review of Federally Funded Studies Related to LEP Students. Final Analytical Report.* Arlington, VA: Development Associates, Inc., 1993.

APPENDICES

Appendix A List of Acronyms for Key Organizations, Groups, and Programs

Appendix B Definitions of Key Terms on the Inclusion of Students with Disabilities and Limited English Proficient Students Assessments

Appendix C Resources for More Information on Inclusion and Assessment Issues

APPENDIX A

List of Acronyms for Key Organizations, Groups, and Programs

For a report such as this one, which provides information on a variety of federal and state programs conducted by a number of different organizations, the widespread use of acronyms may be somewhat confusing to a reader not familiar with them. Even for those experienced in this jargon, the sheer quantity of different (and sometimes quite similar) acronyms may seem a lot like alphabet soup. Thus, in order to help readers of this report, a list of the most frequently used acronyms is provided. The following organizations, groups, and programs are, or have been, involved in some way with the topic of inclusion of students with special needs in assessments, and their acronyms are referenced in various sections of this report.

ADA—Americans with Disabilities Act

AERA—American Educational Research Association

AIR—American Institutes for Research

APA—American Psychological Association

ASAP—Association of State Assessment Programs

CCSSO—Council of Chief State School Officers

CRESST—National Center for Research on Evaluation, Standards, and Student Testing

DID—Division of Innovation and Development

ED—U.S. Education Department

ESEA—Elementary and Secondary Education Act

ESSI—Education Statistics Services Institute

ETS—Educational Testing Service

GWU/CEEE—George Washington University Center for Equity and Excellence in Education

IASA—Improving America's Schools Act

IDEA—Individuals with Disabilities Education Act

IEP—Individualized Education Plan

JCTP—Joint Committee on Testing Practices

LEP—Limited English Proficient (or Proficiency)

NABE—National Association of Bilingual Educators

NAE—National Academy of Education

NAEP—National Assessment of Educational Progress

NAGB—National Assessment Governing Board

NAS—National Academy of Sciences

NASDSE—National Association of State Directors of Special Education

NCBE—National Clearinghouse for Bilingual Education

NCEO—National Center on Educational Outcomes

NCES—National Center for Education Statistics

NCME—National Council on Measurement in Education

NCREL—North Central Regional Education Laboratory

NELS—National Education Longitudinal Study

NRC—National Research Council

NVS—NAEP Validity Studies

OBEMLA—Office of Bilingual Education and Minority Languages Affairs

OCR—Office for Civil Rights

OERI—Office of Educational Research and Improvement

OGC—Office of the General Counsel

OSEP—Office of Special Education Programs

OSERS—Office of Special Education and Rehabilitation Services

OUS—Office of the Under Secretary

PRAEP—Puerto Rico Assessment of Educational Progress

SCASS—State Collaborative on Assessment and Student Standards

SD—Students with Disabilities

SSAP—State Student Assessment Programs

TGPA—Technical Guidelines for Performance Assessments

TSA—(NAEP) Trial State Assessment

APPENDIX B

Definitions of Key Terms on the Inclusion of Students with Disabilities and Limited English Proficient Students Assessments

The following definitions are provided for reference to many of the terms used in the study of the inclusion of students with disabilities or LEP students in assessments. Some of the terminology is based on a set of definitions included in the document, "*Assessment Terminology,*" developed under the guidance of Jim Ysseldyke of NCEO in conjunction with the SCASS group on Special Education Assessment. In addition, terminology for LEP students is based on definitions provided in the "Glossary of Terms and Acronyms" included in the Center for Equity and Excellence in Education's document *Promoting Excellence: Ensuring Success for Limited English Proficient Students* (GWU/CEEE 1996). Please note that this list is not meant to be comprehensive of the terminology used in the field; it is provided as a handy resource for some of the most commonly-used terms included in this report.

Accommodations: Alterations in how a test is presented to the test taker or how the test taker responds; includes a variety of alterations in presentation format, response format, setting in which the test is taken, timing, or scheduling. The alterations do not substantially change level, content, or performance criteria. The changes are made in order to level the playing field, that is, to provide equal opportunity to demonstrate what is known.

Accountability: A systematic method to assure to those inside and outside of the educational system that schools are moving in desired directions—commonly included elements are goals, indicators of progress toward meeting those goals, analysis of data, reporting procedures, and consequences or sanctions. Consequences or sanctions might include additional or fewer resources, removal of accreditation, provision of professional development training, etc.

Adaptations: Changes made in an assessment approaches, usually to the test format, that allow students to participate in the assessment. Adaptations include Braille and large-print versions of test instruments or use of audio cassettes.

Appropriate Assessment: Discussions of "good practice" in the appropriate design and use of assessments of student learning that touch upon (1) the need for multiple measures aligned with curriculum content and instruction; (2) a focus on monitoring student growth over time; and (3) opportunity for students to demonstrate their knowledge in varied ways. For LEP students, relevant issues include: Do the test results accurately reflect a student's knowledge of the content matter, or do they reflect the student's limited proficiency with the language and context of the test items? Alternative but equivalent strategies, including, but not limited to extra time, assessment in the

primary language or language of instruction, can be used to enable LEP students to express their academic content knowledge.

Assessment: The process of collecting data for the purpose of making decisions about individuals, groups, or systems.

Content ESL: A program of instruction that uses ESL techniques to teach academic content areas. (See also English as a Second Language.)

Content Standards: Statements of the subject-specific knowledge and skills that schools are expected to teach and students are expected to learn. They indicate what students should know and be able to do.

English Language Learner (ELL): A term suggested by researchers of the field (Rivera, 1994; August and Hakuta, 1997) as an alternative to "Limited English Proficient" or "language minority" student. ELL refers to students whose first language is not English and encompasses both students who are just beginning to learn English and those who have already developed considerable proficiency.

English as a Second Language (ESL): A specialized program of instruction in which English is used as the language of instruction in order to develop an LEP student's English proficiency to the level of their mainstream peers. This methodology is primarily focused on grammar, vocabulary, and communication, rather than academic content areas. (See also Content-ESL.)

Exclusion from testing: The act of barring someone from participation in an assessment program.

Exemption from testing: The act of releasing someone from a testing requirement to which others are held.

High stakes testing: Assessment that has significant consequences for an individual or school system, e.g., a high school graduation test that is used to determine whether a student receives a diploma is a high stakes test for the student, whereas a test that determines whether a school receives a financial reward or is accredited is a high stakes test for the school.

Individualized Education Program (IEP): A document which reflects the decisions made by the IEP committee during an IEP meeting. Included in this document is a description of the student's performance level and the corresponding goals and objectives to address the areas of need.

IEP Committee: The group which meets to discuss a student's areas of strength and need, and develops an individualized plan for the student's educational program.

Language Minority: In most parts of the United States, this term refers to a native speaker of a language other than English.

Limited English Proficient (LEP): Students whose proficiency in English has not yet developed to the point where they can fully participate in an English-only instructional environment. LEP is the official term found in federal legislation. (See also English Language Learner.)

Modifications: Substantial changes in what a student is expected to learn and/or demonstrate. The changes include changes in instructional level, content, and performance criteria, and may also include changes in test form or format.

Monolingual: Refers to a program that uses one language or a person who speaks only one language. This term is often used to describe regular, mainstream, school programs in English.

Non-English Proficient (NEP): Describes a person who had not yet begun acquiring or who is in the initial stage of learning English.

Opportunity to Learn Standards: Requirements for educational inputs and processes designed to ensure that all students are given the opportunity to achieve the knowledge and skills contained in national, state, district, and/or school content and performance standards.

Out-of-Level Testing: Administration of a test at a level above or below one generally recommended for a student based on his or her grade level or age. Done to accommodate the ability levels of students who are either much above or below the average of students of their age and thus would not demonstrate the entire range of skills they have.

Participation rate: Number of students with disabilities or LEP students taking a test divided by the number of students with disabilities or LEP students, respectively, at the grade level or corresponding age level (for ungraded students) covered by the assessment.

Performance assessment: A task that requires a student to create an answer or a product rather than simply fill in a blank or select a correct answer from a list; the task performed by the student is intended to simulate real life situations.

Performance standards: Benchmarks for how good a student's skills must be in areas aligned with content standards. Typically, performance standards are indices of level of performance.

Reliability: In measurement, the extent to which it is possible to generalize from an observation of a specific behavior observed at a specific time by a specific person to

observations conducted on a similar behavior, at different times, or by different observers.

Testing: The administration of a particular set of questions to an individual or group of individuals for the purpose of obtaining a score.

Translation: An issue in the selection of native language instructional materials or assessment instruments for LEP students, translation becomes most critical in high stakes testing. In the context of assessment, translation strategies may imply that the native language test is derived from the English language test by direct translation from English to the native language or by culturally appropriate adaptation of the language and content from English to the native language. Each of these methods of translation raises its own validity issues. Selection of a method of translation should be based upon the purpose of the assessment. More fundamental than decisions about the methodology of translation is the question of when it is appropriate to use a native language test. Student literacy and prior content area instruction in the native language are minimum indicators that some type of translation may be an appropriate assessment tool.

Validity: The extent to which a test measures what its authors or users claim it measures. Specifically, test validity concerns the appropriateness of the inferences that can be made on the basis of test results.

APPENDIX C

Resources for More Information on Inclusion and Assessment Issues

A wide variety of resources exist for persons interested in obtaining more information on the topic of the inclusion of students with disabilities and LEP students in large-scale assessments. For this report, all of the following resources were used. Many of the persons and organizations involved in inclusion activities and/or large-scale assessments were directly contacted. In addition, an extensive search of an assortment of resources was conducted. Together, the two approaches were used to gather a comprehensive collection of information on the activities and issues that were discussed throughout this report.

Some of the resources and approaches listed below are those traditionally used to search the literature, such as paper-based articles from various publications. Other resources are based on new technology, such as the electronic medium to communicate the latest information with other interested researchers and newsgroups, and on the World Wide Web (WWW).

A. Clearinghouses

A good place to start investigating topics on assessment, education, and the inclusion of students with disabilities and LEP students is at clearinghouses that collect and disseminate this information. These places tend to have an extensive and varied amount of resources gathered into one central place, including reports, documents, and other publications; databases; and references to other places with information on the topic of interest.

ERIC Clearinghouses

The Educational Resources Information Center, (ERIC) is a national information system designed to provide users with ready access to an extensive body of education-related literature. ERIC was established in 1966 and is supported by the U.S. Department of Education, OERI, and the National Library of Education. Information can be accessed through the ERIC database via the Internet, or by printed documents. Other services provided are the ERIC Document Reproduction Service, for full-text copies of research documents; ERIC Digests, two-page syntheses of research; and the ERIC Components, which include clearinghouses, and support components. For general ERIC information, call 1-800-LETERIC or email at acceric@inet.ed.gov. The World Wide Web address for the ERIC homepage is listed in the next section.

The ERIC Clearinghouses collect, abstract, and index education materials for the ERIC database; respond to requests for information in their subject specific areas; and produce special publications on current research, programs, and practices. More than 20 clearinghouses exist within ERIC, covering a wide variety of education-related topics. Those related to the focus of this report are:

- Assessment and Evaluation
- Disabilities and Gifted Education
- Languages and Linguistics

Each of these clearinghouses can be accessed by going through the ERIC homepage on the WWW.

National Clearinghouse for Bilingual Education

The NCBE is funded by OBEMLA to collect, analyze, synthesize, and disseminate information related to the education of linguistically and culturally diverse learners in the United States. It is operated by the George Washington University, Institute for the Study of Language and Education/Center for Education Policy Studies. NCBE provides information through its online services, a toll-free fax service, and a telephone reference and referral service. The NCBE web site is discussed further in Section B.

A special NCBE Internet service, AskNCBE, is a convenient way for email users to request information and order publications from NCBE or to ask for help with NCBE's online services. Contact AskNCBE through email at askncbe@ncbe.gwu.edu.

National Center on Educational Outcomes

The NCEO, established in 1990 by the Office of Special Education Programs, is the only national center focusing its activities on educational outcomes for all students, including students with disabilities. The Center is part of the College of Education and Human Development at the University of Minnesota. NCEO provides states with national leadership in the identification of indicators to monitor educational outcomes for students and works to promote a national discussion of education goals and indicators of educational outcomes that include students with disabilities. The Center also analyzes and summarizes information on students with disabilities in state, national, and international data collection programs, including large-scale assessments.

NCEO publishes a number of reports, including outcome documents, technical reports, synthesis reports, policy directions, and brief reports, on various issues. These publications can be obtained by contacting the NCEO Publications Office at (612) 626-1530.

B. Gophers and World Wide Web sites

Gophers are menu driven systems providing access to a wide variety of information. Usually, Gopher sites can be accessed by typing GOPHER and the gopher address. In recent years, many gophers have been replaced with World Wide Web sites. The WWW is an Internet service that presents information using hypertext. With the appropriate software, such as Netscape or Mosaic, one can see expanded versions of text with different fonts and colors, and graphic images such as maps, charts, and other figures.

Following are some selected Web sites that may include information on inclusion, accommodations, and the assessment of students with disabilities and LEP students. Note that the addresses listed for these sites are up-to-date as of early 1997, but organizations sometimes

change their addresses, so use of a WWW search engine may be required if a site cannot be contacted directly using the address provided below.

U.S. Department of Education
http://www.ed.gov/

This web site contains an extensive amount of information and descriptions of federal and state programs, the organization of the U.S. Department of Education, and various ED publications. It provides important information on programs sponsored by OERI, OBEMLA, and OSERS, as well as links to web pages for the different offices within the department. Many other on-line educational resources on the Internet also are listed and linked to this site. In addition to the Department's own online library, useful information can be found at sites run by clearinghouses, National Centers, Labs, and other organizations funded by the Department. Several National Research and Development Centers have web sites of interest, for example the Center for Research on the Education of Students Placed at Risk (CRESPAR), the National Center for Research on Cultural Diversity and Second Language Learning, and the Center for Research on Education, Diversity, and Excellence (CREDE).

The three main offices in the U.S. Department of Education that are most involved in providing support to research projects and activities related to the inclusion of special needs students in assessments each have their own home page. The OERI home page contains information on the different institutes, programs, and services within the office, including NCES and the National Institute on the Education of At-Risk Students. The OSERS home page provides information on the programs that assist in educating children with special needs, provide for the rehabilitation of youth and adults with disabilities, and support research to improve the lives of individuals with disabilities. The OBEMLA home page provides information on the role of bilingual education, Title VII programs, news and happenings in the office, funding opportunities, and technical assistance services.

National Center for Education Statistics
http://www.ed.gov/NCES/

Although the NCES home page can be accessed by going through the ED site to OERI, the NCES web site can also be accessed directly by using the above address. Within this home page, one can get information on NCES's many programs, data, and surveys; gain quick access to statistics about education; find, view, and download NCES publications; link to the ED home page and other education sites; and search the site for more information. Data on assessment and surveys that include students with disabilities and LEP students can be found here, including information on NAEP.

National Assessment of Educational Progress
http://www.ed.gov/NCES/NAEP/

The NAEP home page resides within the NCES web site. The NAEP site contains information on the program and from the assessments, including recent NAEP reports and data almanacs. Information on a variety of special studies is also provided, including information on some of the research studies being done to increase the inclusion of students with disabilities and LEP students in the assessment.

American Educational Research Association
http://aera.net/

AERA is a professional organization composed of educational researchers in many of the social sciences. The AERA web site provides information of its divisions and Special Interest Groups, descriptions of its publications, listserves, and a valuable file on resources from around the world for educational researchers. Information on specific resources focusing on inclusion and assessment issues can be found at this web site, including information on a new AERA SIG that focuses on research issues, special projects, and activities related to the topic of increasing inclusion of students with disabilities and LEP students in large-scale assessments.

ERIC Clearinghouse on Assessment and Evaluation
http://ericae2.educ.cua.edu/

As discussed earlier (under the section on "Clearinghouses") this site contains an enormous amount of information on measurement and evaluation, complete text of books, essays, and newsletters on the topic, access to places to search ERIC databases, descriptions of major testing projects, materials related to Goals 2000, and links to other sites containing assessment and evaluation information. In addition, a resource called the Test Locator contains descriptions of over 10,000 test instruments to measure a wide array of interests, aptitudes, skills, and academic achievement.

National Center on Educational Outcomes
http://www.coled.umn.edu/NCEO/

The NCEO website was referenced also in the summary above on clearinghouses. This site links users directly to the many reports available from NCEO, as well as to key staff and researchers at NCEO.

National Clearinghouse for Bilingual Education
http://www.ncbe.gwu.edu/

This web site provides a dynamic way of accessing NCBE (details on this clearinghouse were also provided above). The site includes an online library containing full text versions of NCBE publications and other articles, access to NCBE databases, information about and a library of links related to language, culture, and education. NCBE also has a gopher to connect to the organization's information.

National Center for Research on Evaluation, Standards, and Student Testing
http://cresst96.cse.ucla.edu/index.htm

The CRESST web site provides access to many newsletters, technical reports, and other publications, descriptions of videos and handbooks on alternative assessment, and a large database on alternative assessment. Research papers from studies investigating issues related to NAEP and the assessment of LEP students can be found at this site.

Education Week on the Web
http://www.edweek.org/

This web resource, based on the weekly publication, Education Week, provides news, in-depth analysis, and general information of education-related topics. It also contains a database of

old issues of EdWeek Information on a wide variety of topics, including national and state activities on the inclusion of students with disabilities and LEP students in large-scale assessments can be found at this site.

C. Listserves

Listserves provide a medium for discussion among members interested in a specific area(s) of focus. To participate, one must to have an Internet address and must subscribe to the group. Many Listserves exist, covering a wide variety of topics. Members send messages to all other members of the listserve, and anyone may respond to anyone else's message. The intent is to have an open discussion of any topics of interest. The following Listserves may include discussions related to the topic of inclusion of students with disabilities and LEP students in assessments.

AERA

AERA sponsors a variety of Internet Forums for discussions of educational research on listserves. AERA offers two general lists, 12 division lists, a list for graduate students, plus several lists for Special Interest Groups. One may subscribe to any list by following this procedure:

> Address an email letter to LISTSERVE@asu.edu
> Make the single line contents of that letter read as follows:
> SUB <listname> <yourname>

The AERA forums that may include discussions of inclusion issues are:

AERA	American Educational Research Association list of general interest
AERA-B	Curriculum Studies Forum
AERA-C	Learning and Instruction Forum
AERA-D	Measurement and Research Methodology Forum
AERA-H	School Evaluation and Program Development
AERA-K	Teaching and Teacher Education Forum
AERA-L	Politics and Policy in Education

Note: Other AERA divisions also have listserve forums. See AERA for more information.

K-12ASSESS-L

This electronic forum is for school-based personnel, researchers, and others interested in issues pertaining to the assessment of students in grades K-12. K12@ASSESS is a service of the ERIC Clearinghouse on Assessment and Evaluation and the Computer Center of the Catholic University of America in Washington, DC. The goal of this listserve is to provide educators with a fast, convenient, and topical electronic discussion forum focusing on issues related to educational assessment at the elementary and secondary school levels.

Note: Additional information on other web sites and listserves of interest to the educational community can be found in an article in the Winter 1995 issue of *Educational Measurement:*

Issues and Practice, entitled "Assessment and Evaluation on the Internet" by Liselle Drake, Larry Rudner, and Jean Pierce.

D. Newsgroups

Newsgroups communicate electronically with subscribers by periodically transmitting messages through email. To receive these messages, you must have an Internet address.

Edinfo

Edinfo provides frequent updates on educational initiatives, instructional methods, results of reports, and various programs funded by the U.S. Department of Education The ED Initiatives provides weekly looks at progress on the Secretary of Education's priorities. Edinfo is prepared by staff in the U.S. Department of Education.

To subscribe, address an email message to: listproc@inet.ed.gov. Then write SUBSCRIBE EDINFO YOURFIRSTNAME YOURLASTNAME in the message, then send the message.

NCBE Newsline

Sponsored by the National Clearinghouse for Bilingual Education, NCBE Newsline subscribers will receive the latest announcements from OBEMLA. The NCBE Newsline transmits weekly messages providing detailed information on projects, organizations, reports, conferences, educational opportunities, and bilingual education resources.

To subscribe, contact majordomo@cis.ncbe.gwu.edu.

NABE Action Alert

The National Association of Bilingual Educators (NABE) provides an electronic mail distribution read-only list as a convenient and efficient way to electronically distribute copies of its Action Alerts to both its membership and all other interested parties. Action Alerts are not distributed on any fixed schedule, but are issued several times a year. Recipients can not reply or post to this list.

NABE welcomes comments or suggestions sent to NABE@nabe.org.

E. Newsletters

One may subscribe to receive printed newsletters from a variety of sources. These newsletters are issued on various schedules (although most come by mail every month or so).

"NABE NEWS"

The National Association of Bilingual Educators issues a news magazine about educational equity and excellence through bilingual education. NABE NEWS is published by NABE eight times a year. It covers articles on any issue related to bilingual education, including the testing of English language learners and the performance of LEP students on assessments.

"OBEMLA Fax Newsletter"

The OBEMLA Fax Newsletter is transmitted to subscribers by fax or by email on a

regular basis. This newsletter provides updates of Congressional activities, grant opportunities, and other current news items that are the focus of OBEMLA.

To subscribe, contact NCBE by email at askncbe@ncbe.gwu.edu.

"Datalinks"

The National Center on Educational Outcomes began publishing this newsletter in 1996. Datalinks focuses on issues related to the assessment of students with disabilities, highlights of states' projects to increase the participation of students with disabilities, and updates of activities of other groups and organizations involved in this area. Datalinks is intended to be a networking vehicle for those involved in research projects and activities on this topic. For more information, contact Judy Elliott at NCEO via email at ellio015@gold.tc.umn.edu.

"The Exchange"

This newsletter, published by the State Collaborative on Assessment and Student Standards, of the CCSSO, provides updates of the progress and activities of the various consortia. Summaries of a variety of ongoing projects are included in this monthly publication, as well as schedules of SCASS and other important meetings related to assessment and education. For more information, contact Ed Roeber by email at edroeber@aol.com.

"Improving Performance Assessment"

This newsletter is issued by the SCASS Technical Guideline for Performance Assessment Consortium, of the CCSSO. Publication of this newsletter is funded by a grant from OERI. The newsletter focuses on the work of the SCASS TGPA consortium, and provides updates on their projects. Information on issues related to validity, alternative assessments, state assessment practices, use of accommodations, and participation of students with disabilities and LEP students in large-scale assessments have been included in previous issues. For more information, contact Phoebe Winter via email at pwinter123@aol.com.

There are many other sources of information on students with disabilities and LEP issues in large-scale assessments than the ones listed here. For example, it is also possible to gain access to various databases and datasets from certain organizations, such as from NCES or NCBE. The resources listed in this appendix are to help the interested researcher find additional information on the topic. With the advent of networked systems and electronic communications, it is possible to obtain the latest information available from many sources. In particular, using Internet to explore some of the sites listed here can help you link to many other sites and an increasingly vast wealth of information, including the latest reports on the topic. In fact, you may even be reading this report from the WWW.

*U.S. GOVERNMENT PRINTING OFFICE: 1997-437-946/70455

ISBN 0-16-049132-0

Standard Mail (A)

BEST COPY AVAILABLE